VORLESUNG ANALYSIS III, WINTERSEMESTER 2015/16

JOHANNES EBERT

1. Das Lebesgue-Mass

Hauptgegenstand dieser Vorlesung ist es, zu erklären, wie Funktionen $f : \mathbb{R}^d \to \mathbb{R}$ integriert werden können. Anschaulich wurde in der Analysis I das Integral einer Funktion $f : [a, b] \to [0, \infty)$ als Flächeninhalt zwischen der x-Achse und dem Graphen von f erklärt. Wir haben jedoch den Begriff des Flächeninhaltes niemals rigoros entwickelt.

Der Integralbegriff, den wir im folgenden entwickeln werden, basiert darauf, den Begriff des Flächeninhaltes ernst zu nehmen. Das erste größere Ziel dieser Vorlesung ist die Konstruktion des *Lebesgue-Maßes* $\mu(S)$ einer Teilmenge $S \subset \mathbb{R}^d$. Wir beginnen mit einigen Vorbereitungen.

1.1. **Die erweiterten reellen Zahlen.**

Viele Situationen in der Integrationstheorie erfordern einen korrekten Umgang mit dem Symbol ∞. Dafür führt man die erweiterten reellen Zahlen ein.

Definition 1.1. *Die erweiterten reellen Zahlen sind $[-\infty, \infty] := \mathbb{R} \cup \{\pm\infty\}$. Hierbei ist $\pm\infty \notin \mathbb{R}$. Die Ordnungsrelation auf den reellen Zahlen wird durch die Festsetzung*

$$-\infty < x < \infty$$

für alle $x \in \mathbb{R}$ fortgesetzt. Die Addition und Multiplikation lässt sich wie folgt fortsetzen:

- $\infty + x = \infty \forall x \in (-\infty, \infty]$,
- $-\infty + x = -\infty \forall x \in [-\infty, \infty)$,
- $$\pm\infty \cdot x = \begin{cases} \pm\infty & x > 0 \\ 0 & x = 0 \\ \mp\infty & x < 0. \end{cases}$$

Man beachte, dass die Symbole

$$\infty - \infty;\ \frac{x}{\infty};\ \frac{\infty}{x}$$

nicht erklärt sind. Insbesondere ist $[-\infty, \infty]$ weder mit der Addition noch mit der Multiplikation eine abelsche Gruppe, geschweige denn ein Körper. Auch ist es wichtig, zu bemerken, dass die Kürzungsregel

$$x + y = x + z \Rightarrow z = y$$

in $[-\infty, \infty]$ nur dann gilt, wenn $x \neq \pm\infty$ ist.

Wie in der Analysis I erklärt man Supremum und Infimum einer Teilmenge S als kleinste obere beziehungsweise größte untere Schranke. Jede Teilmenge von

---

Date: 2. Februar 2016.

$[-\infty, \infty]$ besitzt Supremum und Infimum. Konvergenz von Folgen wird wie in Analysis I erklärt. Wir notieren kurz die Definition der Konvergenz gegen $\pm\infty$: Eine Folge $(x_n)_n$ in $[-\infty, \infty]$ konvergiert gegen $+\infty$, falls

$$\forall C \in (0, \infty) \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : x_n \geq C$$

(analog für den Grenzwert $-\infty$). Der Satz aus der Analysis I, dass monotone beschränkte Folgen in $\mathbb{R}$ konvergieren, hat folgende Konsequenz: ist $x_n \in [-\infty, \infty]$ und $x_1 \leq x_2 \leq x_3 \leq \ldots$, so konvergiert die Folge gegen $\sup_n x_n \in [-\infty, \infty]$. Für monoton fallende Folgen y_n gilt analog $\lim_n y_n = \inf_n y_n$.

Sei $x_n \in [-\infty, \infty]$ eine Folge. Setze

$$y_n := \sup_{k \geq n} x_k; \; z_n := \inf_{k \geq n} x_k.$$

Dann gilt

$$y_1 \geq y_2 \geq y_3 \ldots; \; z_1 \leq z_2 \leq z_3 \leq \ldots \text{ und } z_n \leq y_n \text{ für alle } n.$$

Deshalb konvergieren die Folgen y_n und z_n. Wir definieren den *Limes superior* als:

$$\limsup_n x_n := \lim_n y_n = \inf_n \sup_{k \geq n} x_k \tag{1.2}$$

und den *Limes inferior* von x_n als

$$\liminf_n x_n := \lim_n z_n = \sup_n \inf_{k \geq n} x_k. \tag{1.3}$$

Weil stets $z_n \leq y_n$ gilt, folgt

$$\liminf_n x_n \leq \limsup_n x_n.$$

Genauer gesagt haben wir folgende Aussagen.

Lemma 1.4. *Sei x_n eine Folge in $[-\infty, \infty]$ und $a \in [-\infty, \infty]$. Dann gilt:*

(1) *$a > \limsup_n x_n$, dann $a > x_n$ für fast alle $n \in \mathbb{N}$,*
(2) *$a < \limsup_n x_n$, dann $a < x_n$ für unendlich viele n mit $a < x_n$,*
(3) *$a > \liminf_n x_n$, dann $a > x_n$ für unendlich viele n,*
(4) *$a < \liminf_n x_n$, dann $a < x_n$ für fast alle n.*

Beweis. Falls $a > \limsup_n x_n = \lim_n \sup_{k \geq n} x_k$, dann gibt es ein n_0 mit $a > \sup_{k \geq n_0} x_k$, also gilt $a > x_k$ für alle $k \geq n_0$, also für fast alle k. Falls $a < \limsup_n x_n = \lim_n \sup_{k \geq n} x_k$, dann gibt es n_0 mit $a < \sup_{k \geq n} x_k$, für alle $n \geq n_0$. Mit anderen: zu jedem noch so großen $n \geq n_0$ gibt es $k \geq n$ mit $a < x_k$. Also $a < x_n$ für unendlich viele n. Die beiden anderen Fälle werden analog gezeigt. $\square$

Die Kontraposition der Aussagen von Lemma 1.4 lauten:

Lemma 1.5. *Sei x_n eine Folge in $[-\infty, \infty]$ und $a \in [-\infty, \infty]$. Dann gilt:*

(1) *falls $a \leq x_n$ für unendlich viele n, so gilt $a \leq \limsup_n x_n$,*
(2) *falls $a \geq x_n$ für fast alle n, so gilt $a \geq \limsup_n x_n$,*
(3) *falls $a \leq x_n$ für fast alle n, so gilt $a \leq \liminf_n x_n$,*
(4) *falls $a \geq x_n$ für unendlich viele n, so gilt $a \geq \liminf_n x_n$.*

Korollar 1.6. *Eine Folge x_n in $[-\infty, \infty]$ ist genau dann konvergent, wenn $\liminf_n x_n = \limsup_n x_n$. In diesem Fall ist $\lim_n x_n = \liminf_n x_n$.*

Beweis. Sei $x_n \to x$ konvergent. Falls $a < \lim_n x_n$, so gilt $a < x_n$ für fast alle n, also nach Lemma 1.5 $a \leq \liminf_n x_n$. Dies gilt *für jedes* a mit $a < \lim_n x_n$, also folgt $\lim_n x_n \leq \liminf_n x_n$. Ist $b > \lim_n x_n$, so ist $b > x_n$ für fast alle n, also nach Lemma 1.5 $b \geq \limsup_n x_n$. Dies gilt für jedes solche b, also $\lim_n x_n \geq \limsup_n x_n$. Zusammengenommen sehen wir

$$\lim_n x_n \leq \liminf_n x_n \leq \limsup_n x_n \leq \lim_n x_n,$$

woraus die Gleichheit aller dieser Zahlen folgt. Falls $\lim_n x_n = \pm\infty$, so muss dieses Argument leicht variiert werden.

Sei umgekehrt $\liminf_n x_n = \limsup_n x_n =: A$ und wir nehmen an, dass $A \in \mathbb{R}$. Für jedes $\epsilon > 0$ ist dann, nach Lemma 1.4,

$$A - \epsilon < x_n < A + \epsilon$$

für fast alle n. Somit konvergiert x_n gegen A. Der Fall $A = \pm\infty$ erfordert wieder einige kleinere Modifikationen. □

Wir notieren noch

$$\liminf_n x_n = -\limsup_n(-x_n); \ \limsup_n x_n = -\liminf_n(-x_n).$$

In $[0, \infty]$ ist jede Reihe konvergent. Sei nämlich $x_n \in [0, \infty]$ eine Folge. Dann definiert man

$$\sum_{n=1}^{\infty} x_n := \sup_{F \subset \mathbb{N} \text{endlich}} \sum_{n \in F} x_n.$$

Sei allgemeiner I eine beliebige Menge, und $x : I \to [0, \infty]$, $i \mapsto x_i$ eine Abbildung. Dann setzt man

$$\sum_{i \in I} x_i := \sup_{F \subset I \text{endlich}} \sum_{i \in F} x_i \in [0, \infty].$$

Man beachte, dass man auf diese Art nur die Summen von Reihen mit nichtnegativen Gliedern nehmen kann.

Satz 1.7 (Doppelreihensatz). *Seien I, J Mengen und $a : I \times J \to [0, \infty]$, $(i, j) \mapsto a_{i,j}$ eine Abbildung. Dann gilt*

$$\sum_{i \in I}(\sum_{j \in J} a_{i,j}) = \sum_{(i,j) \in I \times J} a_{i,j} = \sum_{j \in J}(\sum_{i \in I} a_{i,j}).$$

Beweis. Aus Symmetriegründen muss nur eine der beiden Gleichungen gezeigt werden, und wir zeigen die erste Gleichheit. Zunächst beobachtet man, dass der Satz ganz sicher stimmt, wenn beide Mengen I und J endlich sind (das folgt letztlich aus der Kommutativität der Addition und mittels vollständiger Induktion). Setze $b_i := \sum_{j \in J} a_{i,j}$. Sei $H \subset I \times J$ endlich. Es gibt dann endliche Teilmengen $E \subset I$, $F \subset J$ mit $H \subset E \times F$. Daraus folgt

$$\sum_{(i,j) \in H} a_{i,j} \leq \sum_{(i,j) \in E \times F} a_{i,j} = \sum_{i \in E} \sum_{j \in F} a_{i,j} \leq \sum_{i \in E} \sum_{j \in J} a_{i,j} \leq \sum_{i \in I} \sum_{j \in J} a_{i,j};$$

die Gleichung stimmt, weil E und F endlich sind. Dies gilt für alle endlichen H, und durch Bilden des Supremum folgt also

$$\sum_{(i,j) \in I \times J} a_{i,j} \leq \sum_{i \in I} \sum_{j \in J} a_{i,j} = \sum_{i \in I} b_i.$$

Nun müssen wir noch die Ungleichung $\sum_{(i,j)\in I\times J} a_{i,j} \geq \sum_{i\in I} b_i$ zeigen. Dies ist nur dann nichttrivial, wenn $\sum_{(i,j)\in I\times J} a_{i,j} < \infty$. In diesem Fall gilt auch $b_i < \infty$, denn es ist (offensichtlich!) $b_i \leq \sum_{(i,j)\in I\times J} a_{i,j}$. Also können wir annehmen, dass $b_i < \infty$ für alle i gilt. Sei $E \subset I$ endlich und $\epsilon > 0$. Für jedes $i \in E$ finden wir eine Teilmenge $F_i \subset J$ mit

$$\sum_{j\in F_i} a_{i,j} \geq b_i - \epsilon.$$

Setze $F = \cup_{i\in E} F_i$; dies ist endlich und es folgt, für alle $i \in E$,

$$\sum_{j\in F} a_{i,j} \geq \sum_{j\in F_i} a_{i,j} \geq b_i - \epsilon.$$

Es folgt dann

$$\sum_{i\in E} b_i \leq \sum_{i\in E}(\sum_{j\in F} a_{i,j} + \epsilon) = \sum_{(i,j)\in E\times F} a_{i,j} + \sharp(E)\epsilon \leq \sum_{(i,j)\in I\times J} a_{i,j} + \sharp(E)\epsilon.$$

Nun war ϵ *beliebig*, und durch Grenzübergang $\epsilon \to 0$ erhalten wir

$$\sum_{i\in E} b_i \leq \sum_{(i,j)\in I\times J} a_{i,j}.$$

Weil $E \subset I$ eine beliebige endliche Teilmenge war, folgt

$$\sum_{i\in I} b_i \leq \sum_{(i,j)\in I\times J} a_{i,j},$$

was zu zeigen war. □

Bemerkung 1.8. *In dem obigen Beweis sind einige der einfachen, aber weitverbreiteten Tricks in der Maßtheorie versammelt:*

(1) *Wenn eine Gleichung $A = B$ in der Maßtheorie gezeigt werden soll, so zeige man $A \leq B$ und $B \leq A$.*
(2) *Man verschaffe sich ein ϵ Platz; statt $A \leq B$ zeige man $A \leq B + \epsilon$ für alle positiven ϵ.*

Der Leser tut gut daran, die zu verinnerlichen: wir brauchen beide Tricks dutzendfach.

1.2. **Maßräume.** Wir definieren nun den abstrakten Kontext, in dem die Integrationstheorie formuliert werden wird.

Definition 1.9. *Es sei X eine Menge. Eine* σ-Algebra *auf X ist eine Menge $\mathcal{B}$ von Teilmengen von X mit folgenden Eigenschaften:*

(1) *$\emptyset \in \mathcal{B}$;*
(2) *ist $S \in \mathcal{B}$, so ist das Komplement $S^c := X \setminus S$ ebenfalls in $\mathcal{B}$;*
(3) *sind $S_n \in \mathcal{B}$, $n \in \mathbb{N}$, so ist $\cup_{n=1}^{\infty} S_n \in \mathcal{B}$.*

Natürlich ist die Potenzmenge $\mathcal{P}(X)$, also die Menge aller Teilmengen von X, eine σ-Algebra. Man beachte, dass nur gefordert wird, dass *abzählbare* Vereinigungen von Mengen in $\mathcal{B}$ wieder in $\mathcal{B}$ sind, nicht beliebige Vereinigungen. Dieser Punkt ist zentral!! Einige einfache Beobachtungen: Sei $\mathcal{B}$ eine σ-Algebra auf X.

(1) Sind $S_n \in X$, $n \in \mathbb{N}$, so ist $\cap_{n=1}^{\infty} S_n$ wieder in $\mathcal{B}$. Grund:

$$\cap_{n=1}^{\infty} S_n = ((\cap_{n=1}^{\infty} S_n)^c)^c = (\cup_{n=1}^{\infty} S_n^c)^c.$$

(2) Endliche Vereinigungen von Elementen in $\mathcal{B}$ sind auch in $\mathcal{B}$: man nehme die leere Menge hinzu.
(3) Die Differenz $A \setminus B = A \cap B^c$ ist wieder in $\mathcal{B}$.

Definition 1.10. *Sei X eine Menge und $\mathcal{B}$ eine σ-Algebra auf X. Ein* Maß *ist eine Abbildung $\mu : \mathcal{B} \to [0, \infty]$ mit*
(1) *$\mu(\emptyset) = 0$,*
(2) *Sind die Mengen $S_n \in \mathcal{B}$, $n \in \mathbb{N}$, paarweise disjunkt, so gilt $\mu(\cup_{n=1}^{\infty} S_n) = \sum_{n=1}^{\infty} \mu(S_n)$ (σ-Additivität).*

Ein Tripel $(X, \mathcal{B}, \mu)$, bestehend aus einer Menge X, einer σ-Algebra $\mathcal{B}$ und einem Maß μ heißt Maßraum.

Beispiel 1.11. *Es sei X eine Menge. Sei $\mathcal{B}$ die σ-Algebra* aller *Teilmengen von X und setze*
$$\mu(S) := \sharp S,$$
die Mächtigkeit von S. Dies ist ein Maß, das Zählmaß. *Etwas allgemeiner sei $a : X \to [0, \infty]$. Setze*
$$\mu'(S) := \sum_{s \in S} a(s).$$

Proposition 1.12. *Sei $(X, \mathcal{B}, \mu)$ ein Maßraum und $S_n \in \mathcal{B}$, $n \in \mathbb{N}$. Dann gilt:*
(1) *$S_1 \subset S_2$, dann $\mu(S_1) \leq \mu(S_2)$,*
(2) *$S_1 \subset S_2$ und $\mu(S_2) < \infty$, dann $\mu(S_1) - \mu(S_2) = \mu(S_1 \setminus S_2)$,*
(3) *$\mu(S_1 \cup S_2) + \mu(S_1 \cap S_2) = \mu(S_1) + \mu(S_2)$,*
(4) *$S_1 \subset S_2 \subset \ldots$, dann $\mu(\cup_n S_n) = \lim_n \mu(S_n)$,*
(5) *$S_1 \supset S_2 \supset \ldots$ und $\mu(S_1) < \infty$, dann gilt $\mu(\cap_n S_n) = \lim_n \mu(S_n)$.*

Beweis. (1): $S_2 = S_1 \cup (S_2 \setminus S_1)$ eine disjunkte Vereinigung, und es folgt $\mu(S_2) = \mu(S_1) + \mu(S_2 \setminus S_1) \geq \mu(S_1)$.

(2): Wir schreiben
$$\mu(S_2) = \mu(S_1) + \mu(S_2 \setminus S_1)$$
und dürfen diese Gleichung umstellen, weil alle drei Größen endlich sind.

(3): Wir zerlegen $S_1 \cup S_2 = (S_1 \cap S_2) \cup (S_2 \setminus S_1) \cup (S_1 \setminus S_2)$ disjunkt. Ferner sind $S_2 = (S_1 \cap S_2) \cup (S_2 \setminus S_1)$ beziehungsweise $S_1 = (S_1 \cap S_2) \cup (S_1 \setminus S_2)$ disjunkte Vereinigungen. Es folgt
$$\mu(S_1 \cup S_2) + \mu(S_1 \cap S_2) = \mu(S_1 \cap S_2) + \mu(S_1 \cap S_2) + \mu(S_2 \setminus S_1) + \mu(S_1 \setminus S_2) = \mu(S_1) + \mu(S_2).$$

(4): Setze $T_1 = S_1$ und $T_n := S_n \setminus S_{n-1}$. Dann ist $T_n \cap T_m = \emptyset$ für $m \neq n$ und $\cup_n T_n = \cup_n S_n$. Daraus folgt
$$\mu(\cup_n S_n) = \sum_{k=1}^{\infty} \mu(T_k) = \lim_n \sum_{k=1}^{n} \mu(T_k) = \lim_n \mu(\cup_{k=1}^{n} T_k).$$
Aber es gilt nach Konstruktion $\cup_{k=1}^{n} T_k = S_1 \cup (S_2 \setminus S_1) \cup \ldots \cup (S_n \setminus S_{n-1}) = S_n$.

(5): Wegen Teil (1) gilt $\mu(S_n) < \infty$ für alle n. Setze $U_n := S_1 \setminus S_n$. Dann gilt $U_1 \subset U_2 \subset U_3 \subset \ldots$ und $\cup_{n=1}^{\infty} U_n = S_1 \setminus \cap_{n=1}^{\infty} S_n$. Hieraus folgt, unter Benutzung von (3), dass
$$\mu(S_1 \setminus \cap_{n=1}^{\infty} S_n) = \lim_n \mu(S_1 \setminus S_n).$$
Nun ist $\mu(S_1) < \infty$. Deshalb können wir diese Gleichung wegen (2) als
$$\mu(S_1) - \mu(\cap_{n=1}^{\infty} S_n) = \lim_n \mu(S_1) - \mu(S_n)$$

schreiben. □

1.3. **Das Lebesgue-Maß.** Das einzige Maß, dass direkt angebbar ist, ist das Zählmaß aus Beispiel 1.11. Unser Vorhaben ist die Konstruktion eines sehr viel subtileren Maßes, des *Lebesgue-Maßes*. Die Definition ist schnell gegeben:

Definition 1.13. *Sei $I = [a, b)$ oder $(a, b]$ oder (a, b) oder $[a, b]$, so ist das* Elementarvolumen $|I| := |b - a|$. *Ein* Quader *ist eine Teilmenge $Q \subset \mathbb{R}^d$ der Form $Q = I_1 \times I_2 \times \ldots \times I_d$, wobei I_i ein beschränktes (offenes, kompaktes oder halboffenes) Intervall ist. Das Elementarvolumen von Q ist*

$$|Q| := \prod_i |I_i|.$$

Ausgehend von diesem einfachen Inhaltsbegriff definiert man

Definition 1.14. *Sei $S \subset \mathbb{R}^d$. Das* äußere Maß *von S ist*

$$\mu^*(S) := \inf\{\sum_{j=1}^{\infty} |Q_j| | Q_j \textit{ abgeschlossener Würfel }, S \subset \cup_{j=1}^{\infty} Q_j\}.$$

Man kann in der Definition abgeschlossene Würfel durch offene Würfel ersetzen, wie in Lemma 1.20 gezeigt wird. Ebenso können wir die Definition etwas umformulieren:

$$\mu^*(S) = \inf\{\sum_{i \in I} |Q_i| | I \text{ abzählbar }, Q_j \text{ abgeschlossener Würfel }, S \subset \cup_{i \in I} Q_i\}.$$

Definition 1.15. *Eine Teilmenge $S \subset \mathbb{R}^d$ heißt* Lebesgue-messbar*, falls zu jedem $\epsilon > 0$ eine offene Menge $S \subset U \subset \mathbb{R}^d$ existiert mit $\mu^*(U \setminus S) \leq \epsilon$. Ist S Lebesgue-messbar, so heißt*

$$\mu(S) := \mu^*(S)$$

das Lebesgue-Maß von S.

Soweit die Definitionen. Das erste große Ziel dieser Vorlesung ist der Beweis, dass das Lebesgue-Maß ein Maß ist, also die in den Definitionen 1.9 und 1.10 geforderten Eigenschaften gelten. Genauer gesagt, haben wir folgendes Ergebnis.

Satz 1.16 (Satz über das Lebesguemaß)**.** *Es gilt:*

(1) *Die Menge $\mathcal{L} = \mathcal{L}_{\mathbb{R}^d}$ der Lebesgue-messbaren Teilmengen des $\mathbb{R}^d$ ist eine σ-Algebra.*
(2) *Das Lebesgue-Maß ist ein Maß auf $\mathcal{L}$. Darüber hinaus gilt:*
(3) *Ist $S \in \mathcal{L}$ eine* Nullmenge*, also $\mu(S) = 0$ und ist $T \subset S$, so ist T Lebesgue-messbar (also in $\mathcal{L}$) und $\mu(T) = 0$.*
(4) *Jede offene (und jede abgeschlossene) Teilmenge von $\mathbb{R}^d$ ist Lebesgue-messbar.*
(5) *Für jedes $S \in \mathcal{L}$ gilt*
$$\mu(S) = \inf\{\mu(U) | S \subset U \textit{offen}\}.$$
(6) *Für jedes $S \in \mathcal{L}$ gilt*
$$\mu(S) = \sup\{\mu(K) | K \subset S \textit{kompakt}\}$$
(7) *Ist $S \in \mathcal{L}$ und $x \in \mathbb{R}^d$, so ist $x + S \in \mathcal{L}$ und $\mu(x + S) = \mu(S)$.*
(8) *Ist Q ein Quader, so gilt $\mu(Q) = |Q|$.*

Die Eigenschaften (3), (4), (5) und (6) haben in der allgemeinen Sprache der Maßtheorie spezielle Namen. Die dritte Eigenschaft besagt, dass der Maßraum $(\mathbb{R}^d, \mathcal{L}, \mu)$ *vollständig ist.* Die vierte Eigenschaft verknüpft den Begriff der Messbarkeit mit den topologischen Eigenschaften von $\mathbb{R}^d$ und besagt, dass das Lebesgue-Maß ein *Borel-Maß* ist. Die fünfte Eigenschaft ist als *äußere Regularität* des Lebesgue-Maßes bekannt, die sechste als *innere Regularität.* Die Eigenschaft (7) bezieht sich auf die algebraische Struktur des $\mathbb{R}^d$ und ist als *Translationsinvarianz* bekannt, und (8) schließlich ist eine *Normierung.* Man kann auch zeigen, dass $\mathcal{L}$ und μ durch die obigen Forderungen eindeutig bestimmt sind, aber dieser Aspekt wird in der Vorlesung keine Rolle spielen.

Der Beweis dieses Satzes ist hochgradig nichttrivial, und relativ lang. Das Ergebnis ist das Fundament der Integrationstheorie, und auch für sich genommen sehr interessant.

1.4. **Beweis der Eigenschaften des Lebesgue-Maßes: Vorbereitungen.** Wir nehmen uns nun den Beweis von Satz 1.16 vor. Dieser ist in eine lange Folge von Lemmata gegliedert.

Auf den ersten Teil mögen Teile des Beweises redundant erscheinen: beispielsweise müssen wir separat beweisen, dass offene und dass abgeschlossene Teilmengen messbar sind, obwohl aus den Axiomen für ein Maß folgt, dass nur eine Eigenschaft nachgewiesen werden muss. Dies liegt in der Natur der Sache! Zunächst eine Vereinbarung: *Im ganzen restlichen Kapitel ist eine "offene Menge" stets eine in $\mathbb{R}^d$ offene Teilmenge des $\mathbb{R}^d$.*

Lemma 1.17. *Das äußere Maß ist monoton, d.h. falls $S \subset T \subset \mathbb{R}^d$, so gilt $\mu^*(S) \leq \mu^*(T)$.* □

Lemma 1.18. *Offene Teilmengen des $\mathbb{R}^d$ sind Lebesgue-messbar. Es gilt $\mu(\emptyset) = 0$.*

Beweis. Dies ist nicht mehr als eine bloße Sprachregelung. Die leere Familie von Würfeln überdeckt die leere Menge, und daher gilt $\mu^*(\emptyset) = 0$. Da aber $\emptyset$ offen ist, folgt die Messbarkeit sofort. Ist $U \subset \mathbb{R}^d$ offen, so ist $\mu^*(U - U) = 0 \leq \epsilon$ für alle $\epsilon > 0$, und daher ist U messbar. □

Die Translationsinvarianz können wir direkt zeigen.

Beweis von Satz 1.16 (7). Sei $S \subset \mathbb{R}^d$. Sei $T_x(y) := x + y$ die Translationsabbildung. Wir zeigen zunächst, dass stets

$$\mu^*(T_x(S)) \leq \mu^*(S)$$

gilt. Um dies einzusehen, betrachte man Würfel W_n mit $S \subset \cup_{n=1}^{\infty} W_n$. Dann ist $T_x(W_n)$ ein Würfel, und es gilt $|T_x(W_n)| = |W_n|$. Weil $T_x(S) \subset \cup_{n=1}^{\infty} T_x(W_n)$ folgt (nach der Definition des äußeres Maßes)

$$\mu^*(T_x(S)) \leq \sum_{n=1}^{\infty} |T_x(W_n)| = \sum_{n=1}^{\infty} |W_n|.$$

Also ist $\mu^*(T_x(S))$ eine untere Schranke der Menge $\{\sum_{n=1}^{\infty} |W_n| | S \cup_{n=1}^{\infty} W_n\} \subset [0, \infty]$, woraus folgt

$$\mu^*(T_x(S)) \leq \inf\{\sum_{n=1}^{\infty} |W_n| | S \cup_{n=1}^{\infty} W_n\} =: \mu^*(S).$$

Nun ist $T_{-x} \circ T_x = \mathrm{id}$, also $S = T_{-x}(T_x(S))$, und daher gilt auch
$$\mu^*(S) \leq \mu^*(T_x(S))$$
nach dem eben gesagten. Also folgt, dass sogar
$$\mu^*(S) = \mu^*(T_x(S))$$
für alle S gilt.

Sei nun S messbar und $\epsilon > 0$. Dann wähle eine offene Menge $U \supset S$ mit $\mu^*(U - S) \leq \epsilon$. Es ist $T_x(U)$ offen (denn T_x^{-1} ist stetig) und außerdem
$$\mu^*(T_x(U) - T_x(S)) = \mu^*(T_x(U - S)) = \mu^*(U - S) \leq \epsilon.$$
Da $\epsilon > 0$ beliebig war, ist $T_x(S)$ als messbar erkannt. $\square$

Alle anderen Eigenschaften des Lebesgue-Maßes sind nichttrivial, und wir beginnen mit einer schwachen Version der σ-Additivität.

Satz 1.19 (σ-Subadditivität des äußeren Maßes)**.** *Seien $S_1, S_2, \ldots \subset \mathbb{R}^d$. Dann gilt*
$$\mu^*(\cup_{j=1}^\infty S_j) \leq \sum_{j=1}^\infty \mu^*(S_j).$$

Beweis. Sei $\epsilon > 0$. Wähle Würfel $W_{j,k}$ mit $S_j \subset \cup_{k=1}^\infty W_{j,k}$ und $\mu^*(S_j) \geq \sum_{k=1}^\infty |W_{j,k}| - \frac{\epsilon}{2^j}$. Es ist dann $S := \cup_{j=1}^\infty S_j \subset \cup_{(j,k)\in\mathbb{N}\times\mathbb{N}} W_{j,k}$ und daher nach Definition des äußeren Maßes
$$\mu^*(S) \leq \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} |W_{j,k}| \overset{!}{=} \sum_{j=1}^\infty \sum_{k=1}^\infty |W_{j,k}| \leq \sum_{j=1}^\infty (\mu^*(S_j) + \frac{\epsilon}{2^j}) = \sum_{j=1}^\infty (\mu^*(S_j)) + \epsilon.$$
Die mit Ausrufezeichen markierte Gleichung stimmt wegen des Doppelreihensatzes. Da ϵ beliebig war, folgt die Behauptung. $\square$

Lemma 1.20. *Ist $S \subset \mathbb{R}^d$, so gilt*
$$\mu^*(S) = \inf\{\sum_{j=1}^\infty |P_j| | P_j \textit{ offener Würfel}, S \subset \cup_{j=1}^\infty P_j\}.$$

Beweis. Sei B das Infimum auf der rechten Seite. Ist $S \subset \cup_j P_j$, P_j offener Würfel, so gilt $S \subset \cup_j \bar{P}_j$, und $|\bar{P}_j| = |P_j|$. Also
$$\sum_{j=1}^\infty |P_j| = \sum_{j=1}^\infty |\bar{P}_j| \geq \mu^*(S)$$
und somit (gehe zum Infimum) $B \geq \mu^*(S)$. Nun sei $S \subset \cup_{j=1}^\infty Q_j$, Q_j abgeschlossener Würfel. Sei $\epsilon > 0$. Wähle einen offenen Würfel $P_j \supset Q_j$ mit $|Q_j| \geq |P_j| - \epsilon/2^j$. Dann gilt
$$\sum_{j=1}^\infty |Q_j| \geq \sum_{j=1}^\infty (|P_j| - \frac{\epsilon}{2^j}) = (\sum_{j=1}^\infty |P_j|) - \epsilon \geq B - \epsilon.$$
Also folgt ($\epsilon \to 0$)
$$\sum_{j=1}^\infty |Q_j| \geq B$$
und daher (gehe zum Infimum über) $\mu^*(S) \geq B$. $\square$

Nun können wir die ersten substantiellen Teile von Satz 1.16 beweisen.

Satz 1.21 (Äußere Regularität)**.** *Für jede Teilmenge $S \subset \mathbb{R}^d$ ist*

$$\mu^*(S) = \inf\{\mu^*(U) | S \subset U \text{ offen}\}.$$

Beweis. Aus der Monotonie des äußeren Maßes (Lemma 1.17) folgt $\mu^*(S) \leq \mu^*(U)$, wenn $S \subset U$. Also gilt, durch Übergang zum Infimum, $\mu^*(S) \leq \inf\{\mu^*(U) | S \subset U \text{offen}\}$. Für die umgekehrte Ungleichung wähle man $\epsilon > 0$ beliebig. Wähle offene Würfel W_j mit $S \subset \cup_{j=1}^{\infty} W_j$ und $\sum_{j=1}^{\infty} |W_j| \leq \mu^*(S) + \epsilon$ (1.20!). Dann ist $\cup_{j=1}^{\infty} W_j$ offen und $\mu^*(\cup_{j=1}^{\infty} W_j) \leq \sum_{j=1}^{\infty} |W_j| \leq \mu^*(S) + \epsilon$. Also $\inf\{\mu^*(U) | S \subset U \text{offen}\} \leq \mu^*(S) + \epsilon$ für alle ϵ, und $\epsilon \to 0$ zeigt $\inf\{\mu^*(U) | S \subset U \text{offen}\} \leq \mu^*(S)$. ☐

Lemma 1.22 (Vollständigkeit)**.** *Sei $S \subset \mathbb{R}^d$ messbar mit $\mu(S) = 0$ und $T \subset S$. Dann ist T messbar (und wegen Lemma 1.17 gilt $\mu(T) = 0$).*

Beweis. Sei $\epsilon > 0$ und $U \supset S$ offen mit $\mu^*(U - S) < \epsilon$. Dann ist $\mu^*(U - T) \leq \mu^*(U)$. Wegen Satz 1.19 (der auch für endliche Vereinigungen gilt), folgt $\mu(U) \leq \mu^*(S) + \mu^*(U - S) \leq 0 + \epsilon$. ☐

Satz 1.23. *Seien $S_1, S_2 \ldots$ messbare Teilmengen des $\mathbb{R}^d$. Dann ist $S := \cup_{j=1}^{\infty} S_j$ messbar.*

Beweis. Sei $\epsilon > 0$. Es existieren offene U_n mit $S_n \subset U_n$ und $\mu^*(U_n - S_n) \leq \frac{\epsilon}{2^n}$. Dann ist $U = \cup_{n=1}^{\infty} U_n$ offen, $S \subset U$. Ferner ist $U - S \subset \cup_{n=1}^{\infty} (U_n - S_n)$ und daher

$$\mu^*(S - U) \leq \sum_{n=1}^{\infty} \mu^*(S_n - U_n) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon,$$

was zu zeigen war. ☐

1.5. **Die Normierung des Lebesgue-Maßes.** Bis jetzt haben wir die topologischen Eigenschaften $\mathbb{R}^d$ nur in formaler Weise ausgenutzt. An zwei Stellen müssen wir wesentlich benutzen, dass $\mathbb{R}^d$ ein vollständiger metrischer Raum ist, und beide Male läuft es auf ein Kompaktheitsargument hinaus. Wir wollen zunächst das äußere Maß eines Quaders berechnen, also zeigen, dass $\mu^*(Q) = |Q|$ gilt. Später zeigen wir, dass Quader auch messbar sind, womit der Beweis von Satz 1.16(8) erbracht ist.

Ein Teil des Beweises beruht auf einem rein kombinatorischen Argument. Zunächst einige Notationen. Für $x \in \mathbb{R}^d$ und $l > 0$ sei

$$W_{x,l}^d := [x_1, x_1 + l] \times \ldots \times [x_d, x_d + l]$$

der Würfel mit Kantenlänge l, dessen einer Eckpunkt in x liegt. Ferner sei, für $r > 0$, $r\mathbb{Z}^d \subset \mathbb{R}^d$ die Menge aller rx, mit $x \in \mathbb{Z}^d$. Ist $Q \subset \mathbb{R}^d$ ein Quader und $N \in \mathbb{N}$, so setzen wir

$$J_N(Q) := \{x \in \frac{1}{N}\mathbb{Z}^d | W_{x,\frac{1}{N}}^d \cap \bar{Q} \neq \emptyset\}$$

und

$$I_N := I_N(Q) := \{x \in \frac{1}{N}\mathbb{Z}^d | W_{x,\frac{1}{N}}^d \subset \mathring{Q}\}.$$

Das folgende Lemma zeigt, dass man das Elementarvolumen auf Zählen und einen Grenzübergang zurückführen kann.

Lemma 1.24 (Diskretisierungslemma)**.** *Es sei $Q \subset \mathbb{R}^d$ ein Quader. Dann gilt*

$$\lim_{N \to \infty} \frac{1}{N^d} \sharp I_N(Q) = \lim_{N \to \infty} \frac{1}{N^d} \sharp J_N(Q) = |Q|.$$

Beweis. Wir betrachten zunächst den Fall $d = 1$. Sei $(a, b) \subset Q \subset [a, b]$, $a \leq b$. Es gilt dann

$$b - a - 2 \leq \sharp I_1(Q) \leq \sharp J_N(Q) \leq b - a + 2,$$

wie man durch eine einfache Fallunterscheidung einsieht. Um $I_N(Q)$ und $J_N(Q)$ abzuschätzen, benutzt man, dass

$$\sharp(I_N(Q)) = \sharp(I_1(NQ))$$

und

$$\sharp(J_N(Q)) = \sharp(J_1(NQ))$$

gilt (NQ ist $\{Ny | y \in Q\}$). Also ist

$$N(b-a) - 2 \leq I_1(NQ) = I_N(Q) \leq J_N(Q) = J_1(NQ) \leq N(b-a) + 2.$$

Division durch N und Grenzübergang $N \to \infty$ zeigt die Behauptung. Nun zum Fall $d > 1$. Es sei $Q = Q_1 \times \ldots \times Q_d$ als Produkt von Intervallen geschrieben. Es gilt dann

$$I_N(Q) = I_N(Q_1) \times \ldots \times I_N(Q_d)$$

und

$$J_N(Q) = J_N(Q_1) \times \ldots \times J_N(Q_d)$$

(warum?). Deswegen ist

$$|Q| = \prod_{j=1}^{n} |Q_j| = \prod_{j=1}^{n} \lim_{N\to\infty} \frac{1}{N} \sharp(I_N(Q_j)) = \lim_{N\to\infty} \frac{1}{N^d} \prod_{j=1}^{d} \sharp(I_N(Q_j)) = \lim_{N\to\infty} \frac{1}{N^d} \sharp(I_N(Q))$$

und analog für J_N statt I_N. □

Korollar 1.25. *Seien Q, $Q_1, \ldots, Q_n$ Quader und $Q \subset \cup_{j=1}^{n} Q_j$. Dann gilt $|Q| \leq \sum_{j=1}^{n} |Q_j|$.*

Beweis. Es gilt $J_N(Q) \subset \cup_{j=1}^{n} J_N(Q_j)$ und daher folgt die Ungleichung durch Grenzübergang aus Lemma 1.24. □

Man kann dieses unscheinbare und scheinbar banale Ergebnis auch direkt zeigen; der Beweis ist aber eher mühselig.

Satz 1.26. *Sei $Q \subset \mathbb{R}^d$ ein Quader. Dann gilt $\mu^*(Q) = |Q|$.*

Beweis. Zunächst zeigen wir $\mu^*(Q) \leq |Q|$. Wir überdecken Q durch die abgeschlossenen Würfel $W^d_{x,\frac{1}{N}}$, $x \in J_N$. Es gilt dann, nach Definition des äußeren Maßes und weil $|W^d_{x,\frac{1}{N}}| = \frac{1}{N^d}$,

$$\mu^*(Q) \leq \sum_{x \in J_N} |W^d_{x,\frac{1}{N}}| = (\sharp J_N) \frac{1}{N^d}.$$

Dies gilt für alle N, und Grenzübergang $N \to \infty$ zeigt $\mu^*(Q) \leq |Q|$ (Lemma 1.24).

Für die umgekehrte Abschätzung $\mu^*(Q) \geq |Q|$ nehmen wir zuerst an, dass Q abgeschlossen ist. Sei $Q \subset \cup_j W_j$, W_j offene Würfel. Wir müssen $\sum_{j=1}^{\infty} |W_j| \geq |Q|$ zeigen.

Nun ist Q kompakt, und $Q \subset \cup_j W_j$ eine *offene Überdeckung*. Nach dem Satz von Heine-Borel existiert ein $n \in \mathbb{N}$ so dass $Q \subset \cup_{j=1}^{n} W_j$. Also gilt (Korollar 1.25)

$$|Q| \leq \sum_{j=1}^{n} |W_j| \leq \sum_{j=1}^{\infty} |W_j|.$$

Dies gilt für jede Überdeckung durch offene Würfel, und somit $|Q| \leq \mu^*(Q)$.

Nun sei Q ein beliebiger Quader und $\epsilon > 0$. Es gibt dann einen abgeschlossenen Quader $P \subset Q$ mit $|P| \geq (1-\epsilon)|Q|$. Aus dem oben bewiesenen folgt dann

$$(1-\epsilon)|Q| \leq |P| = \mu^*(P) \leq \mu^*(Q) \leq \mu^*(\bar{Q}) = |\bar{Q}| = |Q|$$

und für $\epsilon \to 0$ die Gleichheit aller dieser Zahlen. □

Es mag reichlich weit hergeholt erscheinen, hier den Satz von Heine-Borel zu zitieren. Tatsächlich ist es notwendig, an dieser Stelle die Vollständigkeit von $\mathbb{R}$ zu benutzen (man versuche, die Theorie für $\mathbb{Q}^d$ anstelle von $\mathbb{R}^d$ zu entwickeln; Satz 1.26 ist dann falsch).

1.6. **Innere Regularität für offene Mengen.** Seien $S_1, S_2 \dots$ messbare Teilmengen des $\mathbb{R}^d$, welche paarweise disjunkt seien. Wir haben bereits gesehen, dass die Vereinigung $S = \cup_{n=1}^\infty S_n$ messbar ist (1.23) und dass $\mu(\cup_{n=1}^\infty S_n) \leq \sum_{n=1}^\infty \mu(S_n)$ gilt (1.19). Wir müssen aber zeigen, dass Gleichheit gilt, und an dieses Ziel können wir uns nur schrittweise herantasten.

In Zukunft fixieren wir auf dem $\mathbb{R}^d$ die ℓ^∞-Norm $\|x\| := \|x\|_{\ell^\infty} := \max\{|x_1|, \dots, |x_d|\}$. Man bemerke, dass metrische Bälle in dieser Norm nichts anderes als Würfel sind.

Definition 1.27. *Seien $S, T \subset \mathbb{R}^d$. Der* Abstand *von S und T ist*

$$\operatorname{dist}(S,T) := \inf\{\|x-y\| | x \in S, y \in T\}.$$

Lemma 1.28. *Sei $A \subset \mathbb{R}^d$ abgeschlossen und $K \subset \mathbb{R}^d$ kompakt. Falls $A \cap K = \emptyset$, so gilt* $\operatorname{dist}(A,K) > 0$.

Beweis. Man nehme an, dass $\operatorname{dist}(A,K) = 0$. Dann gibt es $x_n \in A$ und $y_n \in K$ mit $\|x_n - y_n\| \to \operatorname{dist}(A,K) = 0$. Weil K kompakt ist, gibt es eine Teilfolge x_{n_k}, welche gegen ein $x \in K$ konvergiert. Dann gilt

$$\|y_{n_k} - x\| \leq \|y_{n_k} - x_{n_k}\| + \|x_{n_k} - x\| \to 0,$$

also $y_{n_k} \to x$. Weil A abgeschlossen ist, so muss $x \in A \cap K$ gelten, ein Widerspruch. □

Man kann diesen Satz auch anders beweisen. Variante 2: die Funktion $f : K \to \mathbb{R}$, $f(x) := \operatorname{dist}(\{x\}, A)$ ist stetig (warum?) und nimmt daher ihr Minimum auf K an... Variante 3: Sei $U_x := B_{\operatorname{dist}(x;A)}(x)$, für $x \in K$. Dann ist $\mathcal{U} = (U_x \cap K)_{x \in K}$ eine offene Überdeckung von K. Sei $\delta > 0$ eine Lebesgue-Zahl dieser Überdeckung. Dann gilt $\delta \leq \operatorname{dist}(K,A)$. Variante 3': Wähle eine endliche Teilüberdeckung von $\mathcal{U}$...

Satz 1.29. *Seien $S, T \subset \mathbb{R}^d$ mit* $\operatorname{dist}(S,T) > 0$. *Dann $\mu^*(S \cup T) = \mu^*(S) + \mu^*(T)$.*

Insbesondere gilt dies, wenn S kompakt ist, T abgeschlossen und beide disjunkt sind (Lemma 1.28).

Beweis. Aus Satz 1.19 folgt $\mu^*(S \cup T) \leq \mu^*(S) + \mu^*(T)$. Sei $\epsilon > 0$ und wähle Würfel W_n mit $S \cup T \subset \cup_n W_n$ sowie

$$\sum_n |W_n| \leq \mu^*(S \cup T) + \epsilon.$$

Sei $0 < \delta < \frac{1}{2} \operatorname{dist}(S,T)$. Wir können die Würfel W_n gegebenenfalls unterteilen, so dass alle Teilwürfel Kantenlänge $< \delta$ haben. Ohne Beschränkung der Allgemeinheit seien alle W_n nun so klein. Sei $J \subset \mathbb{N}$ die Menge aller n mit $W_n \cap S \neq \emptyset$ und $I \subset \mathbb{N}$

die Menge aller n mit $W_n \cap T \neq \emptyset$. Dann ist $I \cap J = \emptyset$, und $S \subset \cup_{n \in J} W_n$ sowie $T \subset \cup_{n \in I} W_n$. Es folgt

$$\mu^*(S) + \mu^*(T) \leq \sum_{n \in J} |W_n| + \sum_{n \in I} |W_n| \leq \sum_{n \in \mathbb{N}} |W_n| \leq \mu^*(S \cup T) + \epsilon.$$

$\epsilon \to 0$ zeigt die Behauptung. □

Noch eine Sprechweise: sei $Q_1, Q_2, Q_3, \ldots$ eine Folge von Würfeln. Wir sagen, dass die Würfel paarweise *fast-disjunkt* sind, falls $\mathring{Q}_n \cap \mathring{Q}_m = \emptyset$ für $m \neq n$ gilt.

Satz 1.30. *Jede offene Teilmenge $U \subset \mathbb{R}^d$ ist fast-disjunkte Vereinigung abzählbar vieler abgeschlossener Würfel.*

Beweis. Ein beliebiger Würfel W hat eine der folgenden drei Eigenschaften:

(1) W ist *schlecht*, wenn $W \cap U = \emptyset$.
(2) W ist *gut*, wenn $W \subset U$.
(3) W ist *möglicherweise gut*, wenn $W \cap U \neq \emptyset$, $W \cap U^c \neq \emptyset$.

Wir erinnern uns nun an die Elementarwürfel $W_{x,1/2^n}$ mit Ecke in $x \in \frac{1}{2^n}\mathbb{Z}^d$ und Kantenlänge $1/2^n$. Wir konstruieren nun eine abzählbare Menge $I = I_1 \cup I_2 \cup I_3 \ldots$ von geeigneten abgeschlossenen Würfeln. Sei I_1 die Menge aller guten Würfel der Form $W_{x,1}$, $x \in \mathbb{Z}^d$. Alle möglicherweise guten Würfel werden in 2^d gleich große Teile geteilt. Wir nehmen wieder die guten zu unserer Menge hinzu, und betrachten die möglicherweise guten, und unterteilen wieder. So fortfahrend erhalten wir eine abzählbare Menge von Würfeln, die alle in U enthalten sind und paarweise fast-disjunkt sind.

Formal sei I_1 die Menge aller guten Würfel der Form $W_{x,1}$, $x \in \mathbb{Z}^d$ und I_n sei die Menge aller guten Würfel der Form $W_{x,\frac{1}{2^n}}$, welche nicht in einem der Würfel aus $I_1 \cup \ldots \cup I_{n-1}$ enthalten sind. Sei $I = \cup_{n=1}^{\infty} I_n$. Die Würfel aus I sind fast-disjunkt und alle in U enthalten.

Die Vereinigung ist ganz U: zu jedem $x \in \mathbb{R}^d$ existiert eine Folge $y_n \in \frac{1}{2^n}\mathbb{Z}^d$ mit $x \in W_{y_n,1/2^n}$ für alle n. Für n groß genug ist $W_{y_n,1/2^n} \subset U$, und es gibt ein kleinstes n_0 mit $W_{y_{n_0},1/2^{n_0}} \subset U$. Dieser Würfel ist gut (oder in einem guten Würfel enthalten). □

Korollar 1.31. *Sei $U \subset \mathbb{R}^d$ offen, und sei*

(1) *$\mathcal{K}$ die Menge aller kompakten Mengen $K \subset U$,*
(2) *$\mathcal{W}$ die Menge aller Teilmengen von U, welche endliche Vereinigungen paarweise disjunkter kompakter Würfel sind.*

Dann gilt

$$\mu^*(U) = \sup_{K \in \mathcal{K}} \mu^*(K) = \sup_{K \in \mathcal{W}} \mu^*(K).$$

Beweis. Es ist klar, dass

$$\mu^*(U) \geq \sup_{K \in \mathcal{K}} \mu^*(K) \geq \sup_{K \in \mathcal{W}} \mu^*(K)$$

und wir müssen $\mu^*(U) \leq \sup_{K \in \mathcal{W}} \mu^*(K)$ zeigen. Für die umgekehrte Ungleichung schreiben wir $U = \cup_{n=1}^{\infty} W_n$ als fast-disjunkte Vereinigung von Würfeln, nach Satz 1.30. Sei $\epsilon > 0$. Zu jedem W_n wählen wir einen abgeschlossenen Würfel $Q_n \subset \mathring{W}_n$ mit $\mu^*(Q_n) \geq \mu^*(W_n) - \frac{\epsilon}{2^n}$. Sei

$$K_m := Q_1 \cup Q_2 \cup \ldots \cup Q_m.$$

Die Menge K_m ist kompakt und *disjunkte* Vereinigung von Würfeln. Wegen Lemma 1.28 und Satz 1.29 gilt (Induktion)

$$\mu^*(K_m) = \sum_{n=1}^m \mu^*(Q_n) \geq (\sum_{n=1}^m \mu^*(W_n)) - \epsilon \overset{1.26}{=} (\sum_{n=1}^m |W_n|) - \epsilon$$

Nun treffen wir eine Fallunterscheidung. Ist $\sum_{n=1}^\infty |W_n| < \infty$, so wähle man ein m mit $\sum_{n=1}^m |W_n| \geq \sum_{n=1}^\infty |W_n| - \epsilon$ und folgert:

$$\mu(K_m) \geq (\sum_{n=1}^m |W_n|) - \epsilon \geq \sum_{n=1}^\infty |W_n| - 2\epsilon \geq \mu^*(S) - 2\epsilon$$

(die letzte Ungleichung ist wieder einfach die Definition des äußeren Maßes). Falls $\sum_{n=1}^\infty |W_n| = \infty$, so wird $\sum_{n=1}^m |W_n|$ beliebig groß. Also können wir in diesem Fall kompakte Teilmengen von U mit beliebig großem äußeren Maß finden. □

Bemerkung: in der Situation des vorherigen Beweises gilt sogar $\mu(U) = \sum_{n=1}^\infty |W_n|$. Dies folgt entweder aus dem Beweis oder aus den Eigenschaften des Lebesgue-Maßes.

1.7. **Konstruktion des Lebesgue-Maßes: Schluss des Beweises.** Wir wollen nun den Beweis von Satz 1.16 zu Ende führen. Zunächst listen wir auf, was noch zu zeigen ist:

- Komplemente messbarer Mengen sind messbar.
- Innere Regularität für beliebige (nicht nur offene) Mengen.
- Das Lebesgue-Maß ist σ-additiv.

Satz 1.32. *Abgeschlossene Teilmengen $A \subset \mathbb{R}^d$ sind messbar.*

Beweis. Sei $A \subset \mathbb{R}^d$ abgeschlossen. Wenn wir zeigen können, dass jede der kompakten Mengen $A_n := A \cap [-n, n]^d$ messbar ist, haben wir gewonnen, wegen 1.23 und weil $A = \cup_{n=1}^\infty A_n$. Also dürfen wir annehmen, dass A kompakt ist. Dann ist A in einem Quader enthalten und hat endliches äußeres Maß.

Sei $\epsilon > 0$. Wegen der schon gezeigten äußeren Regularität (Satz 1.21) gibt es eine offene Menge $A \subset U$ mit $\mu^*(U) \leq \mu^*(A) + \epsilon/2$. Wir wollen $\mu^*(U - A) \leq \epsilon$ zeigen. Auf jeden Fall ist $\mu^*(U - A) \leq \mu^*(U) < \infty$. Wähle $K \subset U - A$ kompakt mit $\mu^*(K) \geq \mu^*(U - A) - \epsilon/2$ (Korollar 1.31). Wegen Satz 1.29 ist

$$\mu^*(A) + \mu^*(K) = \mu^*(A \cup K)$$

und somit

$$\mu^*(A) + \mu^*(U - A) - \epsilon/2 \leq \mu^*(A) + \mu^*(K) = \mu^*(A \cup K) \leq \mu^*(U) \leq \mu^*(A) + \epsilon/2.$$

beziehungsweise $\mu^*(U - A) \leq \epsilon$ (hier nutzt man, dass $\mu^*(A) < \infty$). □

Satz 1.33. *Sei $S \subset \mathbb{R}^d$ messbar. Dann ist S^c messbar.*

Beweis. Zu jedem $n \in \mathbb{N}$ wähle man ein offenes $U_n \supset S$ mit $\mu^*(U_n - S) \leq \frac{1}{n}$. Dann ist U_n^c abgeschlossen und folglich messbar (wegen 1.32). Dann ist auch $T := \cup_{n=1}^\infty U_n^c$ messbar. Aber

$$T = (\cap_{n=1}^\infty U_n)^c \subset S^c$$

und

$$S^c - T = S^c \cap T^c = T^c - S = (\cap_{n=1}^\infty U_n) - S.$$

Also gilt, für alle n,

$$S^c - T \subset U_n - S$$

und daher

$$\mu^*(S^c - T) \leq \frac{1}{n}.$$

Also ist $S^c - T$ eine Nullmenge, und

$$S^c = (S^c - T) \cup T$$

ist als Vereinigung der messbaren Menge T und einer Nullmenge erkannt. Also S^c messbar. □

Die Drehscheibe des gesamten Argumentes ist die innere Regularität des Lebesgue-Maßes.

Lemma 1.34. *Sei $\epsilon > 0$. Dann existieren disjunkte kompakte Mengen $K_1, K_2, \ldots$ in $\mathbb{R}^d$, so dass $\mu^*(R) \leq \epsilon$ gilt, wobei $R := \mathbb{R}^d \setminus (\cup_n K_n)$.*

Beweis. Sei $Q_n = [-n, n]^d$. Nun ist Q_n messbar, und wir finden eine offene Menge $Q_n \subset V_n$ mit $\mu^*(V_n - Q_n) \leq \frac{\epsilon}{2^n}$. Wir können V_n verkleinern, so dass $V_n \subset [-n - 1/2, n+1/2]^d$. Sei $V_0 = \emptyset$. Nun setze $K_n := Q_n - V_{n-1}$. Die Menge K_n ist kompakt; es gilt $K_n \cap K_m = \emptyset$ für $n \neq m$. Außerdem ist nach Konstruktion $R = \mathbb{R}^d \setminus \cup_{n=1}^\infty K_n = \cup_{n=1}^\infty (V_n - Q_n)$. Wegen der Subadditivität gilt somit

$$\mu^*(R) \leq \sum_{n=1}^\infty \mu^*(V_n - Q_n) \leq \epsilon.$$

□

Satz 1.35. *Für $S \subset \mathbb{R}^d$ messbar gilt $\sup_{K \subset S kompakt} \mu(K) = \mu(S)$.*

Beweis. Sei S messbar. Dann ist S^c messbar. Wähle $U \supset S^c$ offen mit $\mu^*(U - S^c) \leq \epsilon$ und setze $A := U^c$. Dann ist $A \subset S$ und $U - S^c = U \cap S = S - U^c = S - A$, d.h. $\mu^*(S - A) \leq \epsilon$. Somit

$$\mu(S) \leq \mu(A) + \mu(S - A) \leq \mu(A) + \epsilon.$$

Soweit haben wir gezeigt, dass $\sup_{A \subset S \text{abgeschlossen}} \mu^*(A) = \mu^*(S)$.

Seien nun K_n und R wie in Lemma 1.34 und $A_n := K_n \cap A$ (kompakt!). Es gilt

$$\mu^*(A) \leq \mu^*(A \cap R) + \mu^*(\cup_n A_n) \leq \epsilon + \mu^*(\cup_n A_n) \overset{1.19}{\leq} \epsilon + \sum_n \mu^*(A_n).$$

Sezue $L_m := A_1 \cup \ldots \cup A_m$. Dann ist L_m kompakt, und $\mu(L_m) = \sum_{n=1}^m \mu(A_n)$, wegen 1.29. Fallunterscheidung: ist $\sum_n \mu^*(A_n) < \infty$, so wähle m mit $\sum_{n=1}^m \mu(A_n) \geq \sum_{n=1}^\infty \mu(A_n) - \epsilon$. Dann gilt

$$\mu(A) \leq 2\epsilon + \sum_{n=1}^m \mu^*(A_n) = \mu(L_m) + 2\epsilon.$$

Falls $\sum_n \mu^*(A_n) = \infty$, so wird das Maß von L_m beliebig groß. □

Satz 1.36. *Das Lebesgue-Maß ist σ-additiv.*

Beweis. Seien zunächst S_1, S_2 messbar und disjunkt. Wir behaupten $\mu(S_1 \cup S_2) = \mu(S_1) + \mu(S_2)$. Hiervon ist die Ungleichung $\leq$ schon bewiesen. Sei $\epsilon > 0$. Falls S_1 oder S_2 unendliches Maß haben, so auch $S_1 \cup S_2$ (Monotonie, Lemma 1.17) und das ganze ist trivial. Also nehmen wir $\mu(S_i) < \infty$ an. Wegen der inneren Regularität gibt es kompakte $K_i \subset S_i$ mit $\mu(S_i) \leq \mu(K_i) + \epsilon/2$. Da $K_1 \cap K_2 = \emptyset$ folgt

$$\mu(S_1) + \mu(S_2) \leq \mu(K_1) + \mu(K_2) + \epsilon = \mu(K_1 \cup K_2) + \epsilon \leq \mu(S_1 \cup S_2) + \epsilon.$$

Dann $\epsilon \to 0$.

Durch Induktion folgt $\mu(\cup_{j=1}^n S_j) = \sum_{j=1}^n \mu(S_j)$, falls die S_j paarweise disjunkt sind. Ist $S = \cup_{j=1}^\infty S_j$ eine abzählbare disjunkte Vereinigung, so ist

$$\sum_{j=1}^n \mu(S_j) = \mu(\cup_{j=1}^n S_j) \leq \mu(S).$$

Grenzübergang $n \to \infty$ zeigt dann $\sum_{j=1}^\infty \mu(S_j) \leq \mu(S)$, und die andere Ungleichung ist schon gezeigt worden. □

1.8. **Nicht Lebesgue-messbare Teilmengen.** Man wird nach einigem Herumprobieren feststellen, dass man keine Mengen angeben kann, welche nicht Lebesgue-messbar sind. Dennoch ist nicht jede Teilmenge des $\mathbb{R}^d$ Lebesgue-messbar, wie wir jetzt sehen werden.

Satz 1.37 (Vitali)**.** *Jedes nichtleere offene Intervall $I \subset \mathbb{R}$ enthält eine Teilmenge, welche nicht Lebesgue-messbar ist.*

Der Beweis des Satzes beruht auf dem Auswahlaxiom der Mengenlehre:

Auswahlaxiom. *Ist $f : X \to Y$ eine surjektive Abbildung von Mengen, dann gibt es eine Abbildung $g : Y \to X$ mit $f \circ g = \mathrm{id}$.*

Beweis von Satz 1.37. Wir führen auf $\mathbb{R}$ folgende Äquivalenzrelation ein:

$$x \sim y \Leftrightarrow x - y \in \mathbb{Q}.$$

Die Menge der Äquivalenzklassen sei mit $\mathbb{R}/\mathbb{Q}$ bezeichnet. Sei $q : \mathbb{R} \to \mathbb{R}/\mathbb{Q}$ die Restklassenabbildung. Diese ist surjektiv. Wir können folgende Verfeinerung vornehmen: Sei $I \subset \mathbb{R}$ ein nichtleeres offenes Intervall. Dann existiert zu jedem $x \in \mathbb{R}$ ein $y \in \mathbb{Q}$, so dass $x + y \in I$. Daraus folgt, dass die Einschränkung

$$q|_I : I \to \mathbb{R}/\mathbb{Q}$$

ebenfalls surjektiv ist.

Das Auswahlaxiom besagt nun, dass eine Abbildung $r : \mathbb{R}/\mathbb{Q} \to I$ existiert mit $q \circ r = \mathrm{id}$. Sei $S \subset I$ das Bild von r. Wir behaupten, *dass S nicht messbar ist.* Dies beruht auf drei Beobachtungen:

(1) Es gilt
$$S = \{x \in \mathbb{R} | x = r(q(x))\},$$
(2) für $x \in \mathbb{Q}$, $x \neq 0$ ist $S \cap (x + S) = \emptyset$,
(3) $\cup_{x \in \mathbb{Q}}(x + S) = \mathbb{R}$.

Warum ist das richtig?

(1) Falls $x = r(q(x))$, so ist $x \in S$. Ist $x \in S$, dann gibt es ein $z \in \mathbb{R}/\mathbb{Q}$ mit $x = r(z)$. Weil $q \circ r = \mathrm{id}$, ist dann $q(x) = q(r(z)) = z$, d.h. $x = r(z) = r(q(x))$.
(2) Sei $y \in S \cap (x+S)$. Dann gilt $y, y-x \in S$. Weil $x \in \mathbb{Q}$ folgt $q(y-x) = q(y)$. Es folgt aus (1): $y - x = r(q(y-x)) = r(q(y)) = y$, also $x = 0$.

(3) Für beliebiges $x \in \mathbb{R}$ ist $x \sim r(q(x))$, also $y := x - r(q(x)) \in \mathbb{Q}$. Aber $r(q(x)) \in S$, also $x \in y + S$.

Wir nehmen nun an, dass S messbar ist. Es gilt, wegen der σ-Additivität und Translationsinvarianz des Lebesgue-Maßes (beachte, dass $\mathbb{Q}$ abzählbar ist):

$$\infty = \mu(\mathbb{R}) = \mu(\coprod_{x \in \mathbb{Q}} x + S) = \sum_{x \in \mathbb{Q}} \mu(x + S) = \sum_{x \in \mathbb{Q}} \mu(S).$$

Somit muss $\mu(S) > 0$ sein, denn andernfalls wäre die Summe 0. Sei nun $I = (a, b)$ und $J = (a - 1, b + 1)$. Dann gilt: $|x| < 1 \Rightarrow x + S \subset J$. Also $\cup_{x \in \mathbb{Q}, |x|<1} x + S \subset J$, also

$$\sum_{x \in \mathbb{Q}, |x|<1} \mu(S) = \sum_{x \in \mathbb{Q}, |x|<1} \mu(x + S) \leq \mu(J).$$

Das ist ein Widerspruch, weil es unendlich viele $x \in \mathbb{Q}$ mit $|x| < 1$ gibt. $\square$

Aus dem Beweis können wir noch weitere Schlüsse ziehen:

Korollar 1.38. *Sei $S \subset (0, 1) \subset \mathbb{R}$ wie im Beweis von Satz 1.37 konstruiert. Sei $T \subset S$ Lebesgue-messbar. Dann gilt $\mu(T) = 0$.*

Beweis. Weil

$$\cup_{x \in \mathbb{Q}, |x|<1} x + S \subset (-1, 2)$$

eine disjunkte Vereinigung ist, folgt

$$\cup_{x \in \mathbb{Q}, |x|<1} x + T \subset (-1, 2),$$

und aus der σ-Additivität und Translationsinvarianz des Lebesgue-Maßes folgt

$$\sum_{x \in \mathbb{Q}, |x|<1} \mu(T) = \sum_{x \in \mathbb{Q}, |x|<1} \mu(x + T) = \mu(\cup_{x \in \mathbb{Q}, |x|<1} x + T) \leq 3$$

und das kann nur stimmen, wenn $\mu(T) = 0$. $\square$

Korollar 1.39. *Jede Lebesgue-messbare Teilmenge $T \subset \mathbb{R}$ mit $\mu(T) > 0$ enthält eine nicht messbare Menge $S \subset T$.*

Beweis. Wir zeigen, dass eine messbare Menge T, deren Teilmengen alle messbar sind, eine Nullmenge sein muss. Sei wieder S eine Vitali-Menge wie im Beweis von Satz 1.37. Angenommen, $T \cap (x + S)$ sei messbar für alle $x \in \mathbb{Q}$. Weil $T \cap (x + S) = x + ((-x + T) \cap S)$, sehen wir, dass $(-x + T) \cap S$ für alle $x \in \mathbb{Q}$ messbar ist. Dann muss, nach dem vorherigen Korollar, $(-x + T) \cap S$ eine Nullmenge sein. Somit ist $T \cap (x + S)$ eine Nullmenge, und nach der σ-Additivität auch $T = \cup_{x \in \mathbb{Q}} T \cap (x + S)$. $\square$

1.9. **Borelmengen.** Bevor wir zum Integral kommen, sind noch einige etwas abstrakte Formalitäten zu erledigen. Die folgenden Dinge sind von rein technischen Interesse.

Lemma 1.40. *Es sei X eine Menge und $\mathcal{A} \subset \mathcal{P}(X)$ eine Menge von Teilmengen von X. Dann gibt es genau eine σ-Algebra $\mathcal{B}(\mathcal{A})$ auf X mit folgenden Eigenschaften:*

(1) $\mathcal{A} \subset \mathcal{B}(\mathcal{A})$.
(2) *Ist $\mathcal{B}$ eine weitere σ-Algebra auf X mit $\mathcal{A} \subset \mathcal{B}$, dann gilt $\mathcal{B}(\mathcal{A}) \subset \mathcal{B}$.*

$\mathcal{B}(\mathcal{A})$ heißt die von $\mathcal{A}$ erzeugte σ-Algebra.

Das Lemma hat gewisse formale Ähnlichkeit mit der folgenden Aussage, welche aus der Linearen Algebra bekannt ist. Ist V ein Vektorraum und $S \subset V$ eine Teilmenge, so gibt es genau einen Unterraum $W \subset V$, so dass $S \subset W$ und so dass für jeden anderen Unterraum $U \subset V$ mit $S \subset U$ gilt: $W \subset U$. Natürlich ist W die Menge aller Linearkombinationen von Vektoren aus S, kann aber auch als der Durchschnitt aller Unterräume $U \subset V$ mit $S \subset U$ konstruiert werden.

Beweis. Die Eindeutigkeit ist klar: haben $\mathcal{B}(\mathcal{A})$ und $\mathcal{B}'(\mathcal{A})$ die geforderten Eigenschaften, so gilt $\mathcal{B}(\mathcal{A}) \subset \mathcal{B}'(\mathcal{A}) \subset \mathcal{B}(\mathcal{A})$. Es gibt auf jeden Fall eine σ-Algebra, welche $\mathcal{A}$ enthält, nämlich die Potenzmenge $\mathcal{P}(X)$. Wir wollen $\mathcal{B}(\mathcal{A})$ als Durchschnitt aller σ-Algebren, welche $\mathcal{A}$ enthalten, konstruieren. Formal geht das wie folgt: Sei $\mathbf{I} \subset \mathcal{P}(\mathcal{P}(X)))$ die Menge aller σ-Algebren $\mathcal{B}$ mit $\mathcal{A} \subset \mathcal{B}$. Weil $\mathcal{P}(X) \in \mathbf{I}$, ist $\mathbf{I} \neq \emptyset$. Nun setzen wir

$$\mathcal{B}(\mathcal{A}) := \cap_{\mathcal{B} \in \mathbf{I}} \mathcal{B}.$$

Einige sehr abstrakte, aber letzlich triviale Überlegungen zeigen, dass $\mathcal{B}(\mathcal{A})$ tatsächlich die geforderten Eigenschaften hat. □

Definition 1.41. *Sei $X = \mathbb{R}^d$ oder allgemeiner ein metrischer Raum. Sei $\mathcal{U} \subset \mathcal{P}(X)$ die Menge aller offenen Teilmengen von X. Dann heißt $\mathcal{B}_X := \mathcal{B}(\mathcal{U})$ die* Borel-σ-Algebra. *Die Elemente von $\mathcal{B}(\mathcal{U})$ heißen* Borel-Mengen. *Dies ist die kleinste σ-Algebra auf X, welche alle offenen Mengen enthält.*

Satz 1.42. *Sei $\mathcal{L}_{\mathbb{R}^d}$ die σ-Algebra aller Lebesgue-messbaren Teilmengen von $\mathbb{R}^d$. Dann gilt $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{L}_{\mathbb{R}^d}$. Sei $\mathcal{W} \subset \mathcal{P}(X)$ die Menge aller offenen Würfel in $\mathbb{R}^d$. Dann gilt $\mathcal{B}_{\mathbb{R}^d} = \mathcal{B}(\mathcal{W})$, das heißt, die Borel-$\sigma$-Algebra auf $\mathbb{R}^d$ wird von den offenen Würfeln erzeugt.*

Man kann zeigen, dass $\mathcal{B}_{\mathbb{R}^d} \neq \mathcal{L}_{\mathbb{R}^d}$ gilt, mit anderen Worten: es gibt Lebesgue-messbare Teilmengen, welche keine Borelmengen sind (wenn $\mathcal{B}_{\mathbb{R}^d} = \mathcal{L}_{\mathbb{R}^d}$ stimmen würde, könnten wir uns diese Formalitäten sparen). Der Unterschied liegt in den Nullmengen, spielt aber leider in späteren Beweisen (Substitutionsformel in höheren Dimensionen) eine leidige Rolle.

Beweis. Wir haben gezeigt, dass $\mathcal{U}$, die Menge aller offenen Teilmengen in $\mathbb{R}^d$, in $\mathcal{L}$ liegt ("offene Mengen sind Lebesgue-messbar“). Also gilt nach Lemma 1.40 $\mathcal{B}_{\mathbb{R}^d} := \mathcal{B}(\mathcal{U}) \subset \mathcal{L}$.

Für die zweite Aussage reicht es, zu beweisen, dass $\mathcal{U} \subset \mathcal{B}(\mathcal{W})$, denn dann folgt $\mathcal{B}(\mathcal{U}) \subset \mathcal{B}(\mathcal{W})$, und $\mathcal{B}(\mathcal{W}) \subset \mathcal{B}(\mathcal{U})$ ist ohnehin klar. Sei nun $U \in \mathcal{U}$, also $U \subset \mathbb{R}^d$ offen. Schreibe $U = \cup_{n=1}^{\infty} W_n$ als abzählbare Vereinigung von abgeschlossenen Würfeln. Nun existiert, weil U offen ist, zu jedem W_n ein offener Würfel V_n mit $W_n \subset V_n \subset U$. Es gilt dann $U = \cup_{n=1}^{\infty} V_n$, also $U \in \mathcal{B}(\mathcal{W})$. □

Satz 1.43. *Die σ-Algebra $\mathcal{L}_{\mathbb{R}^d}$ wird von den offenen Mengen und den Nullmengen erzeugt.*

Beweis. Sei $S \subset \mathbb{R}^d$ Lebesgue-messbar. Dann gibt es offene Mengen $U_n \supset S$ so dass $\mu(U_n - S) \leq \frac{1}{n}$. Dann ist $\cap_n U_n \setminus S$ eine Nullmenge, und $S = (\cap_n U_n) \setminus (\cap_n U_n \setminus S)$ liegt in der σ-Algebra, welche von den offenen Mengen und den Nullmenge erzeugt wird. □

2. Das Lebesgue-Integral

Bis jetzt haben wir die Konstruktion des Lebsgue-Maßes gesehen. Wir wenden uns nun wieder allgemeinen Maßräumen $(X, \mathcal{B}, \mu)$ zu und wollen Funktionen $X \to \mathbb{R}^n$ bzw. $X \to [0, \infty]$ integrieren. Der erste Schritt besteht darin, eine geeignete Klasse von Funktionen zu finden, welche überhaupt als Integranden in Frage kommen.

2.1. **Messbare Funktionen.**

Definition 2.1. *Ein* Messraum *ist ein Paar* $(X, \mathcal{B})$*, bestehend aus einer Menge* X *und einer* σ*-Algebra* $\mathcal{B}$ *auf* X*.*

Definition 2.2. *Seien* $(X, \mathcal{B})$ *und* $(Y, \mathcal{C})$ *Messräume und* $f : X \to Y$ *eine Abbildung. Dann heißt* f messbar*, falls gilt: ist* $S \in \mathcal{C}$*, so ist* $f^{-1}(S) \in \mathcal{B}$*.*

Es sei hier daran erinnert, dass eine Abbildung $f : X \to Y$ metrischer Räume genau dann stetig ist, wenn $f^{-1}(U)$ offen in X ist, für jede offene Teilmenge $U \subset Y$. Man beachte die formale Analogie. Wie stetige Abbildungen komponieren sich auch messbare Abbildungen sehr einfach:

Lemma 2.3. *Seien* $(X, \mathcal{B}), (Y, \mathcal{C})$ *und* $(Z, \mathcal{D})$ *Messräume und* $(X, \mathcal{B}) \xrightarrow{g} (Y, \mathcal{C}) \xrightarrow{f} (Z, \mathcal{D})$ *messbare Abbildungen. Dann ist die Komposition* $f \circ g$ *messbar.*

Beweis. Klar: $S \in \mathcal{D}$, dann $f^{-1}(S) \in \mathcal{C}$, dann $g^{-1}(f^{-1}(S)) \in \mathcal{B}$. Aber $g^{-1}(f^{-1}(S)) = (f \circ g)^{-1}(S)$. □

Normalerweise ist eine σ-Algebra unvorstellbar groß, und wir brauchen ein überprüfbares Kriterium, um Messbarkeit nachzuweisen.

Lemma 2.4. *Sei* $(X, \mathcal{B})$ *ein Messraum,* Y *eine Menge und* $\mathcal{A} \subset \mathcal{P}(Y)$ *sowie* $f : X \to Y$ *eine Abbildung. Dann sind äquivalent:*
1. $f : (X, \mathcal{B}) \to (Y, \mathcal{B}(\mathcal{A}))$ *ist messbar.*
2. *Für jede Menge* $S \in \mathcal{A}$ *ist* $f^{-1}(S) \in \mathcal{B}$*.*

Beweis. Weil $\mathcal{A} \subset \mathcal{B}(\mathcal{A})$, ist die Schlussfolgerung $1 \Rightarrow 2$ offensichtlich. Für $2 \Rightarrow 1$ betrachten wir die Menge $\mathcal{C} \subset \mathcal{P}(Y)$ aller Teilmengen $S \subset Y$, so dass $f^{-1}(S) \in \mathcal{B}$. Vorausgesetzt ist $\mathcal{A} \subset \mathcal{C}$ und zu zeigen ist $\mathcal{B}(\mathcal{A}) \subset \mathcal{C}$. Nach Lemma 1.40 genügt es, zu zeigen, dass $\mathcal{C}$ eine σ-Algebra auf Y ist. Dies ist aber klar: erstens ist $f^{-1}(\emptyset) = \emptyset \in \mathcal{B}$, also $\emptyset \in \mathcal{C}$. Ist zweitens $S \in \mathcal{C}$, so gilt $f^{-1}(S^c) = (f^{-1}(S))^c$, und weil $f^{-1}(S) \in \mathcal{B}$, so ist auch $(f^{-1}(S))^c \in \mathcal{B}$, also $S^c \in \mathcal{C}$. Sind drittens $S_n \in \mathcal{C}$, $n \in \mathbb{N}$, so ist $f^{-1}(\cup_n S_n) = \cup_n f^{-1}(S_n)$, und weil $f^{-1}(S_n) \in \mathcal{B}$ und $\mathcal{B}$ eine σ-Algebra ist, so liegt auch $f^{-1}(\cup_n S_n)$ in $\mathcal{B}$, also $\cup_n S_n \in \mathcal{C}$. □

Lemma 2.5. *Es sei* $f : \mathbb{R}^d \to \mathbb{R}^n$ *stetig. Dann sind* $f : (\mathbb{R}^d, \mathcal{L}_{\mathbb{R}^d}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ *und* $f : (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ *messbar.*

Im allgemeinen ist $f : (\mathbb{R}^d, \mathcal{L}_{\mathbb{R}^d}) \to (\mathbb{R}^n, \mathcal{L}_{\mathbb{R}^n})$ nicht messbar, auch wenn f stetig ist.

Beweis. Sei $U \subset \mathbb{R}^n$ offen. Dann ist $f^{-1}(U) \subset \mathbb{R}^d$ offen, weil f stetig ist. Somit ist $f^{-1}(U) \in \mathcal{B}_{\mathbb{R}^d} \subset \mathcal{L}_{\mathbb{R}^d}$. Weil $\mathcal{B}_{\mathbb{R}^n} = \mathcal{B}(\mathcal{U})$, erledigt Lemma 2.4 den Rest geräuschlos. □

Lemma 2.6. *Sei* $(X, \mathcal{B})$ *ein Messraum und* $f : X \to \mathbb{R}^n$*. Dann sind äquivalent:*

(1) $f : (X, \mathcal{B}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ *ist messbar.*
(2) *Jede Komponentenfunktion* $f_i : (X, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ *ist messbar* $(i = 1, \ldots, n)$.

Beweis. $1 \Rightarrow 2$. Die Koordinatenprojektion $p_i : \mathbb{R}^n \to \mathbb{R}$, $p_i(x) = x_i$, ist stetig. Nun ist $f_i = p_i \circ f$, und betrachte

$$(X, \mathcal{B}) \xrightarrow{f} (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) \xrightarrow{p_i} (\mathbb{R}, \mathcal{B}_{\mathbb{R}}).$$

Nach Lemma 2.5 und Lemma 2.3 ist also f_i messbar.

$2 \Rightarrow 1$. Seien $f_1, \ldots, f_n : (X, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ messbar. Sei $W = (a_1, b_1) \times \ldots \times (a_n, b_n)$ ein offener Würfel. Dann gilt

$$f^{-1}(W) = \cap_{k=1}^n f_k^{-1}(a_k, b_k);$$

somit ist $f^{-1}(W)$ Schnitt messbarer Mengen, also selber messbar. Weil die Borel-σ-Algebra $\mathcal{B}_{\mathbb{R}^n}$ von den offenen Würfeln erzeugt wird (Lemma 1.42), ist also f messbar, nach Lemma 2.4. □

Satz 2.7. *Sei* $(X, \mathcal{B})$ *ein Messraum und seien* $f, g : (X, \mathcal{B}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ *messbar und* $a \in \mathbb{R}$. *Dann gilt*
(1) $f + g$ *ist messbar.*
(2) $\|f\|$ *ist messbar.*
(3) af *ist messbar.*
(4) *Für* $n = 1$ *sind* $f_+ := \max\{0, f\} = \frac{1}{2}(f + |f|)$ *und* $f_- = \min\{0, -f\} = \frac{1}{2}(|f| - f)$ *messbar.*
(5) *Falls* $n = 1$ *oder das Ziel* $\mathbb{C}$ *ist, so ist* fg *messbar.*

Beweis. Das geht leicht von der Hand. Zum Beispiel ist $f + g$ die Komposition

$$(X, \mathcal{B}) \xrightarrow{h} (\mathbb{R}^{2n}, \mathcal{B}_{\mathbb{R}^{2n}}) \xrightarrow{k} (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}),$$

mit $h(x) = (f(x), g(x)$, sowie $k(y, z) = y + z$. h ist messbar (nutze Lemma 2.6) und k ist stetig, also $f + g$ messbar. Analog erledigt man 5, und 2,3,4 ist ähnlich, aber einfacher. □

Wir treffen nun folgende

Vereinbarung 1. *Sei* $(X, \mathcal{B})$ *ein Messraum. Wenn wir im folgenden sagen, dass* $f : X \to \mathbb{R}^n$ *messbar ist, dann meinen wir, dass* $f : (X, \mathcal{B}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ *messbar ist.*

Wir wollen auch Funktionen $f : X \to [-\infty, \infty]$ betrachten.

Definition 2.8. *Die Borel-σ-Algebra* $\mathcal{B}_{[-\infty,\infty]}$ *auf* $[-\infty, \infty]$ *ist die von* $\{(a, \infty] | a \in \mathbb{R}\}$ *erzeugte σ-Algebra.*

Es gilt

$$[-\infty, c] = ((c, \infty])^c \in \mathcal{B}_{[-\infty,\infty]};$$

$$[-\infty, b) = \cup_{n=1}^{\infty} [-\infty, b - \frac{1}{n}] \in \mathcal{B}_{[-\infty,\infty]}$$

sowie

$$(a, b) = (a, \infty] \cap [-\infty, b) \in \mathcal{B}_{[-\infty,\infty]}.$$

Deshalb liegen alle offenen Intervalle in $\mathcal{B}_{[-\infty,\infty]}$, und somit alle Borelmengen $S \subset \mathbb{R}$. Wir schließen hieraus:

Satz 2.9. *Eine Funktion $f : X \to [-\infty, \infty]$ ist genau dann messbar, wenn die Teilmengen $f^{-1}(a, \infty])$ für jedes $a \in \mathbb{R}$ messbar sind.*

Nun sei X ein Messraum und $f_n : X \to [-\infty, \infty]$ eine Folge messbarer Funktionen. Wir bilden aus dieser Folge neue Funktionen, nämlich

$$\sup_n f_n, \inf_n f_n, \liminf_n f_n, \limsup_n f_n.$$

Ganz präzise ist $(\sup_n f_n)(x) := \sup_n(f_n(x))$ gemeint (und analog in den anderen Fällen).

Satz 2.10. *Seien $f_n : X \to [-\infty, \infty]$ messbar, $n \in \mathbb{N}$. Dann sind $\sup_n f_n$, $\inf_n f_n$ messbar. Ferner sind $\liminf_n f_n$ und $\limsup_n f_n$ messbar. Insbesondere sind punktweise Grenzwerte messbarer Funktionen wieder messbar.*

Beweis. Sei $f := \sup_n f_n$ und $a \in \mathbb{R}$. Dann ist $f^{-1}((a, \infty]) = \cup_n f_n^{-1}((a, \infty])$ messbar, und also ist f messbar nach 2.9. Die Messbarkeit von $\inf_n f_n$ folgt analog (beachte $\inf_n f_n = -\sup_n(-f_n)$). Außerdem beachte man

$$\limsup_n f_n(x) = \inf_{n \in \mathbb{N}} \sup_{k \geq n} f_k(x); \ \liminf_n f_n(x) = \sup_{n \in \mathbb{N}} \inf_{k \geq n} f_k(x).$$

□

Definition 2.11. *Sei $(X, \mathcal{B}, \mu)$ ein Maßraum. Eine messbare Funktion $f : X \to \mathbb{R}^n$ heißt* Stufenfunktion, *falls $f(X) \subset \mathbb{R}^n$ eine endliche Menge ist, d.h. falls f nur endlich viele Werte annimmt.*

Eine Stufenfunktion f heißt Treppenfunktion, *falls $\mu(\{x \in X | f(x) \neq 0\}) < \infty$.*

Wir definieren die *charakteristische Funktion* $\chi_S : X \to \mathbb{R}$ einer Menge $S \subset X$ durch

$$\chi_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S. \end{cases}$$

Man bemerke, dass χ_S genau dann messbar ist, wenn S eine messbare Menge ist.

Satz 2.12. *Sei $f : X \to [0, \infty]$ messbar. Dann gibt es eine Folge von Stufenfunktionen $f_n : X \to \mathbb{R}$, so dass*

$$0 \leq f_1(x) \leq f_2(x) \leq \dots$$

und $\lim_n f_n(x) = f(x)$.

Sei $f : X \to \mathbb{R}^n$ messbar. Dann gibt es eine Folge von Stufenfunktionen $f_n : X \to \mathbb{R}^d$ mit $|f_n(x)| \leq |f(x)|$ für alle x und $\lim_n f_n(x) = f(x)$.

Beweis. Im ersten Fall setze

$$A_{n,k} = \{x \in X | \frac{k}{2^n} < f(x) \leq \frac{k+1}{2^n}$$

für $k \in \mathbb{N}$. Setze

$$f_n(x) := \sum_{k=1}^{4^n - 1} \frac{k}{2^n} \chi_{A_{n,k}}.$$

Dies tut es, wie man direkt nachsieht (Bild).

Im zweiten Falle betrachte zuerst $n = 1$. Schreibe $f = f_+ - f_-$ mit nichtnegativen Funktionen. Aus dem ersten Teil nimmt man sich die Funktionenfolgen und betrachtet die Differenz. Im allgemeinen Fall $n > 1$ gehe man komponentenweise vor. □

Sei $(X, \mathcal{B}, \mu)$ ein Maßraum und $S \in \mathcal{B}$, also S sei eine messbare Teilmenge von X. Wir definieren die Einschränkung der σ-Algebra $\mathcal{B}$ auf S. Setze

$$\mathcal{B}|_S := \{T \subset S | T \in \mathcal{B}\}.$$

Also besteht $\mathcal{B}|_S$ aus allen Teilmengen von S, welche in X messbar sind. Man überlegt sich leicht, dass

$$\mathcal{B}|_S = \{T \cap S | T \in \mathcal{B}\}$$

gilt: denn ist $T \in \mathcal{B}$, so gilt $T \cap S \in \mathcal{B}$ (denn S und T sind beide messbar) und daher $S \cap T \in \mathcal{B}|_S$. Ist umgekehrt $T \in \mathcal{B}|_S$, so können wir $T = T \cap S$ als Schnitt zweier Mengen in $\mathcal{B}$ schreiben. Nun sei μ ein Maß auf $\mathcal{B}$. Wir setzen dann

$$\mu|_S(T) := \mu(T),$$

falls T eine messbare Teilmenge von S ist. Dass $\mu|_S$ ein Maß ist, ist offensichtlich.

Lemma 2.13. *Sei $(Y, \mathcal{A})$ ein weiterer Messraum. Falls $f : (X, \mathcal{B}) \to (Y, \mathcal{A})$ messbar ist, so ist die Einschränkung $f|_S : (S, \mathcal{B}|_S) \to (Y, \mathcal{A})$ messbar. Ist umgekehrt $g : (S, \mathcal{B}|_S) \to (Y, \mathcal{A})$ messbar und $y_0 \in Y$ ein fest gewählter Punkt, so ist die durch*

$$\tilde{g}(x) := \begin{cases} g(x) & x \in S \\ y_0 & x \notin S \end{cases}$$

definierte Fortsetzung von g eine messbare Abbildung $(X, \mathcal{B}) \to (Y, \mathcal{A})$.

In der Praxis ist der Fall $Y = \mathbb{R}^n$ oder $[-\infty, \infty]$ relevant. In diesem Falle wird man $y_0 = 0$ wählen.

Beweis. Dies ist eine einfache Konsequenz aus den Definitionen. □

2.2. **Das Integral nichtnegativer Funktionen.** Sei von nun an $(X, \mathcal{B}, \mu)$ ein Maßraum. Wir wollen messbare Funktionen $f : X \to \mathbb{R}^n$ integrieren. Dabei geht man in drei Schritten vor:

(1) Integriere Stufenfunktionen $f : X \to [0, \infty)$.
(2) Integriere messbare Funktionen $f : X \to [0, \infty]$. In diesen beiden Fällen kann das Integral den Wert $+\infty$ annehmen.
(3) Erkläre die Integrierbarkeit messbarer Funktionen $f : X \to \mathbb{R}^n$ und definiere das Integral.

Sei $f : X \to [0, \infty)$ eine Stufenfunktion. Sei $a_0 = 0$ und seien $a_1, \dots, a_n$ die von Null verschiedenen Werte von f und $A_k := f^{-1}(a_k)$. Setze

$$\int_X f(x) d\mu(x) := \sum_{k=1}^n a_k \mu(A_k) \in [0, \infty].$$

Für eine messbare Teilmenge $S \subset X$ setze

$$\int_S f(x) d\mu(x) = \int_X \chi_S f d\mu(x).$$

Die Notation $d\mu(x)$ ist gewählt, damit sowohl das Maß μ als auch die Integrationsvariable klar sind. Wenn kein Zweifel über x und μ besteht oder über X, so schreiben wir oft einfach

$$\int_X f d\mu = \int_X f = \int f.$$

Lemma 2.14. *Sei $(X,\mathcal{B},\mu)$ ein Maßraum. Seien $f,g : X \to [0,\infty]$ Stufenfunktionen. Dann gilt:*

(1) *Die Abbildung $\nu : \mathcal{B} \to [0,\infty]$, $\nu(S) := \int_S f d\mu$ ist ein Maß. Ist $\mu(S) = 0$, so folgt $\nu(S) = 0$.*
(2) $\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu$.
(3) *Für $c \geq 0$ gilt $\int_X c f d\mu = c \int_X f d\mu$.*

Beweis. 1. Es ist klar, dass $\nu(\emptyset) = 0$, und es ist auch klar, dass $\int_S f = 0$, wenn $\mu(S) = 0$. Seien $S_1, S_2, \dots$ messbar und disjunkt, $S = \cup_j S_j$. Dann gilt

$$\int_S f d\mu = \sum_{k=1}^n a_k \mu(S \cap A_k) = \sum_{k=1}^n \sum_{j=1}^\infty a_k \mu(S_j \cap A_k) = \sum_{j=1}^\infty \sum_{k=1}^n a_k \mu(S_j \cap A_k) = \sum_{j=1}^\infty \int_{S_j} f d\mu.$$

Die zweite Gleichung ist die σ-Additivität von μ und die dritte der Doppelreihensatz.

2. Seien nun $b_0 = 0$, $b_1, \dots, b_m$ die von Null verschiedenen Werte von g und $B_j := g^{-1}(b_j)$. Nach Teil 1 ist

$$\int_X (f+g) d\mu = \sum_{j=1}^m \sum_{k=1}^n \int_{A_k \cap B_j} (f+g) d\mu = \sum_{j=1}^m \sum_{k=1}^n (a_k + b_j) \mu(A_k \cap B_j).$$

Andererseits folgt

$$\int_X f d\mu + \int_X g d\mu = \sum_{k=1}^n a_k \mu(A_k) + \sum_{j=1}^m b_j \mu(B_j) = \sum_{k=1}^n \sum_{j=1}^m a_k \mu(A_k \cap B_j) + \sum_{j=1}^m \sum_{k=1}^n b_j \mu(B_j \cap A_k).$$

3. ist klar. □

Korollar 2.15. *Sei X ein Maßraum und $f,g : X \to [0,\infty)$ Stufenfunktionen. Dann gilt*

(1) *Falls $S_1 \subset S_2 \subset S_3 \subset \dots$ eine aufsteigende Folge messbarer Teilmengen von X ist, und $S = \cup_{n \geq 1} S_n$, so gilt $\int_S f(x) d\mu(x) = \lim_{n\to\infty} \int_{S_n} f(x) d\mu(x)$.*
(2) *Falls $f \leq g$, so gilt $\int_X f d\mu \leq \int_X g d\mu$.*
(3) *Falls $S \subset T$ messbar ist, so gilt $\int_S f d\mu \leq \int_T f d\mu$.*

Beweis. 3. folgt sofort daraus, dass $\nu(S) := \int_S f d\mu$ ein Maß ist, und 1. ebenfalls (wegen Proposition 1.12). Für 2. notiert man, dass $g = f + h$ gilt, wobei $h : X \to [0,\infty)$ eine Stufenfunktion ist. Es gilt dann

$$\int_X g d\mu = \int_X f d\mu + \int_X h d\mu \geq \int_X f d\mu.$$

□

Definition 2.16. *Sei $(X,\mathcal{B},\mu)$ ein Maßraum und $f : X \to [0,\infty]$ messbar. Wir definieren*

$$\int_X f(x) d\mu := \sup\{\int_X g d\mu | 0 \leq g \leq f \textit{Stufenfunktion}\} \in [0,\infty].$$

Für eine messbare Teilmenge $S \subset X$ setzen wir

$$\int_S f(x) d\mu(x) = \int_X \chi_S(x) f(x) d\mu(x).$$

Lemma 2.17. *Sei $(X, \mathcal{B}, \mu)$ ein Maßraum und $f, g : X \to [0, \infty]$ messbar sowie $c \in [0, \infty)$. Dann gilt*

(1) $f \leq g$, *dann* $\int_X f \leq \int_X g$.
(2) $S \subset T \subset X$ *messbar, dann* $\int_S f \leq \int_T f$.
(3) $\int_X cf(x)d\mu(x) = c \int_X f(x)d\mu(x)$.
(4) *Falls* $\mu(S) = 0$, *so ist* $\int_S f d\mu = 0$.

Beweis. 1. Sei $h \leq f$ eine Stufenfunktion. Dann gilt $\int_X h \leq \int_X g$, weil $h \leq g$, nach der Definition des Integrals. Diese Ungleichung gilt für alle solchen Stufenfunktionen, und Übergang zum Supremum zeigt $\int_X f \leq \int_X g$.

2. Das folgt sofort aus 1, weil $f\chi_S \leq f\chi_T$.

3. Für $c = 0$ ist die Aussage banal, also sei $c > 0$. Dann

$$\int_X cf = \sup\{\int_X g | 0 \leq g \leq cf\} = \sup\{\int_X ch | 0 \leq ch \leq cf\} = c \sup\{\int_X h | 0 \leq ch \leq cf\} = c \int_X f,$$

wobei das Supremum stets über die Menge aller Stufenfunktionen mit den angegebenen Eigenschaften zu nehmen ist. Die dritte Gleichung nutzt $\int_X cg = c \int_X g$ für Stufenfunktionen.

4. Ist klar, weil die entsprechende Aussage für Stufenfunktionen gilt. □

Wir kommen nun zu dem ersten wirklich interessanten Satz über das Integral, dem Satz über monotone Konvergenz. Dieser Satz bringt die ganze Integrationstheorie ins Laufen. Die σ-Additivität geht ganz entscheidend ein, in Gestalt von Lemma 2.14.

Satz 2.18 (Satz von Beppo-Levi)**.** *Sei $f_n : X \to [0, \infty]$ messbar. Es gelte $0 \leq f_1(x) \leq f_2(x) \leq \ldots$ für alle x, und $f(x) := \lim_n f_n(x)$ sei die Grenzfunktion. Dann gilt*

$$\lim_n \int_X f_n(x) d\mu = \int_X \lim_n f_n(x) d\mu.$$

Beweis. Wir beweisen wieder zwei Ungleichungen, und folgen bei der schwierigen Ungleichung der "ein ϵ an Raum“-Strategie.

Es gilt $f_n \leq f$, und also $\int_X f_n \leq \int_X f$, wegen der Monotonie des Integrals. Weil dies für alle n gilt, folgt

$$A := \lim_n \int_X f_n \leq \int_X f.$$

Sei $0 \leq g \leq f$ eine Stufenfunktion und $0 < \epsilon < 1$. Setze

$$S_n := \{x \in X | f_n(x) \geq (1 - \epsilon)g(x)\}.$$

Dies ist eine messbare Menge (g ist endlich, also ist $f_n - (1 - \epsilon)g$ wohldefiniert und messbar). Es gilt stets

$$S_n \subset S_{n+1},$$

weil die Funktionenfolge f_n monoton wächst.

Ferner ist $\cup_n S_n = X$. Grund: ist $g(x) = 0$, so gilt offenbar $x \in S_n$ für alle n. Ist $g(x) > 0$, dann ist $f(x) > (1 - \epsilon)g(x)$, und also $f_n(x) \geq (1 - \epsilon)g(x)$ für großes n, somit $x \in S_n$ für n groß.

Weil $S \mapsto \int_S (1 - \epsilon)g d\mu$ nach Lemma 2.14 ein Maß ist, folgt

$$(1 - \epsilon) \int_X g d\mu = \int_X (1 - \epsilon) g d\mu \overset{2.15}{=} \lim_n \int_{S_n} (1 - \epsilon) g d\mu.$$

Aber
$$\int_{S_n} (1-\epsilon) g d\mu \leq \int_{S_n} f_n d\mu \leq \int_X f_n d\mu.$$
Zusammengenommen folgt
$$(1-\epsilon)\int_X g d\mu \leq \lim_n \int_X f_n d\mu = A.$$
Dies gilt für alle $\epsilon \in (0,1)$; daher ergibt sich
$$\int_X g d\mu \leq A.$$
Weil das für alle Stufenfunktionen $g \leq f$ gilt, folgt durch Übergang zum Supremum
$$\int_X f d\mu \leq A,$$
was zu zeigen war. □

Wir kommen zu zwei wichtigen Folgerungen.

Lemma 2.19. *Seien $f, g : X \to [0,\infty]$ messbar. Dann gilt*
$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

Beweis. Wähle monoton wachsende Folgen h_n, k_n von Stufenfunktionen, welche punktweise gegen f bzw. g konvergieren. Nach Beppo-Levi gilt dann
$$\int_X f d\mu + \int_X g d\mu = \lim_n \int_X h_n d\mu + \lim_n \int_X k_n d\mu = \lim_n \int_X (h_n + k_n) d\mu = \int_X (f+g) d\mu.$$
Hier haben wir genutzt, dass das Integral von Stufenfunktionen additiv ist. □

Lemma 2.20 (Fatou's Lemma)**.** *Seien $f_n : X \to [0,\infty]$ messbar. Dann gilt*
$$\int_X \liminf_n f_n(x) d\mu \leq \liminf_n \int_X f_n(x) d\mu.$$

Hier kann eine strikte Ungleichung auftauchen, auch dann, wenn sowohl f_n punktweise konvergiert, als auch $\int_X f_n(x) d\mu(x)$ konvergiert. Folgendes Beispiel illustriert dies. Sei $X = \mathbb{R}$ mit dem Lebesgue-Maß. Betrachte die Funktionenfolge $f_n(x) := \frac{1}{n}\chi_{[-n^2,n^2]}(x)$. Dann konvergiert f_n punktweise (sogar gleichmäßig) gegen die Nullfunktion. Also gilt $\int_{\mathbb{R}} \liminf_n f_n(x) d\mu = \int_X 0 d\mu = 0$. Andererseits ist $\int_{\mathbb{R}} f_n(x) d\mu(x) = \frac{1}{n}\mu([-n^2, n^2]) = \frac{2n^2}{n} = 2n$, und daher $\lim_n \int_{\mathbb{R}} f_n(x) d\mu(x) = +\infty$.

Dieses Beispiel ist aus mancherlei Gründen nützlich. Erstens ist es hilfreich, um sich die Richtung der Ungleichung in Lemma 2.20 zu merken. Hier ist ein anderes Beispiel, welches zeigt, dass der Satz von Beppo-Levi für monoton *fallende* Funktionenfolgen *falsch* ist. Setze $f_n(x) := \frac{1}{n}$. Dann gilt $\lim_n f_n(x) = 0$ für alle x, aber $\int_{\mathbb{R}} f_n = \infty$ für alle n.

Beweis. Sei $g_n(x) := \inf_{k \geq n} f_k(x)$. Dann gilt, für alle $k \geq n$:
$$g_n \leq f_k \Rightarrow \int_X g_n \leq \int_X f_k.$$
Also auch
$$\int_X g_n \leq \inf_{k \geq n} \int_X f_k.$$

Es ist $g_n(x) \leq g_{n+1}(x) \leq \ldots$ und konvergiert gegen $\liminf_n f_n(x)$. Nach Beppo-Levi ist dann

$$\int_X \liminf_n f_n(x) d\mu = \lim_n \int_X g_n(x) d\mu.$$

Ferner gilt

$$\lim_n \inf_{k \geq n} \int_X f_k = \liminf_n \int_X f_n.$$

Alles in allem erhalten wir $\int_X \liminf_n f_n(x) d\mu \leq \liminf_n \int_X f_n(x) d\mu$, wie behauptet. □

2.3. **Das Integral vektorwertiger Funktionen.** Das nächste Ziel ist es nun, Funktionen $X \to \mathbb{R}^n$ zu integrieren. Während wir messbaren Funktionen $f : X \to [0, \infty]$ stets ein Integral $\int_X f \in [0, \infty]$ zuweisen können, braucht man hier eine Bedingung, die der absoluten Konvergenz ähnelt. Wir bezeichnen die ℓ^2-Norm auf $\mathbb{R}^n$ mit $x \mapsto |x|$; die Doppelstriche wird anderen Normen auf Vektorräumen integrierbarer Funktionen vorbehalten sein.

Definition 2.21. *Sei $f : X \to \mathbb{R}^n$ messbar. Wir sagen, dass f* integrierbar *ist, wenn*

$$\int_X |f(x)| dx < \infty.$$

Die Menge aller integrierbaren Funktionen wird mit $\mathscr{L}^1(X, \mathbb{R}^n)$ bezeichnet.

Was ist das Integral? Zunächst betrachte $n = 1$. Wir können f in zwei Funktionen zerlegen, nämlich als

$$f(x) = f_+(x) - f_-(x), f_\pm \geq 0.$$

Zur Erinnerung: $f_+(x) := \max(f(x), 0)$ und $f_-(x) := \min(-f(x), 0)$. Weil $|f_\pm| \leq |f|$, gilt $\int_X f_\pm < \infty$, und wir können

$$\int_X f d\mu := \int_X f_+ d\mu - \int_X f_- d\mu \in \mathbb{R}$$

zumindest definieren. Nun gehen wir zu vektorwertigen Funktionen über. Ist $f = (f_1, \ldots, f_n)$ in $\mathscr{L}^1(X; \mathbb{R}^n)$, so ist $|f_i| \leq |f|$, damit f_i in $\mathscr{L}^1(X; \mathbb{R})$ und

$$\int_X f d\mu := \sum_{i=1}^k (\int_X f_i d\mu) e_i$$

eine sinnvolle Definition. Hierbei sind $e_i \in \mathbb{R}^n$ die Standardbasisvektoren.

Satz 2.22.
1. *$\mathscr{L}^1(X, \mathbb{R}^n)$ ist ein Vektorraum.*
2. *Die Abbildung $\int_X : \mathscr{L}^1(X, \mathbb{R}^n) \to \mathbb{R}^n$, $f \mapsto \int_X f d\mu$, ist linear.*
3. *Ist $A : \mathbb{R}^n \to \mathbb{R}^m$ linear, so gilt $A(\int_X f) = \int_X A(f)$.*
4. *Ferner gilt*

$$|\int_X f d\mu| \leq \int_X |f(x)| d\mu.$$

Beweis. Sind $a \in \mathbb{R}$, $f, g \in \mathscr{L}^1(X; \mathbb{R}^n)$, so ist

$$\int |af(x) + g(x)| d\mu(x) \leq \int |a||f(x)| d\mu(x) + \int |g(x)| d\mu(x) < \infty,$$

und somit ist $af + g \in \mathscr{L}^1(X;\mathbb{R}^n)$, womit die erste Aussage auch schon gezeigt ist. Für die Linearität nehmen wir uns zuerst den Fall $n = 1$ vor. Sei $a \in \mathbb{R}$ und $f \in \mathscr{L}^1(X;\mathbb{R})$. Für $a \geq 0$ gilt

$$(af)_+ = af_+;\ (af)_- = af_- \Rightarrow \int af = \int af_+ - \int af_- = a(\int f_+ - \int f_-) = a \int f;$$

und für $a \leq 0$ ist

$$(af)_+ = -af_-;\ (af)_- = -af_+ \Rightarrow \int af = \int -af_- + \int af_+ = a(\int f_+ - \int f_-) = a \int f.$$

Falls $f, g \in \mathscr{L}^1$, so sei $h = f + g$. Es gilt dann

$$h_+ - h_- = f_+ - f_- + g_+ - g_-$$

beziehungsweise

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

Also gilt

$$\int h_+ + \int f_- + \int g_- = \int h_- + \int f_+ + \int g_+,$$

und weil alle Integrale endlich sind, können wir umstellen:

$$\int h_+ - \int h_- = + \int f_+ - \int f_- + \int g_+ - \int g_-,$$

und das war zu zeigen. Die Linearität der Abbildung $\int_X : \mathscr{L}^1(X;\mathbb{R}^n) \to \mathbb{R}^n$ folgt sofort aus dem Fall $n = 1$ und der Definition.

Für Teil (3) sei $A = (a_{ij})$ eine Matrix. Dann ist

$$Af(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} f_j(x) e_i,$$

also

$$\int Af = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} (\int f_j(x)) e_i = A(\int f).$$

Für die Dreiecksungleichung erinnern wir zunächst an folgende Konsequenz der Cauchy-Schwarz-Ungleichung: ist $y \in \mathbb{R}^n$, so gilt

$$|y| = \sup_{|v| \leq 1} \langle y, v \rangle.$$

Es gilt dann, für f integrierbar,

$$|\int_X f(x)| = \sup_{|v|\leq 1} \langle \int_X f(x) d\mu, v \rangle = \sup_{|v|\leq 1} \int_X \langle f(x) d\mu, v \rangle \leq \sup_{|v|\leq 1} \int_X |f(x)||v| d\mu = \int_X |f(x)| d\mu.$$

Die erste Gleichung ist die Konsequenz aus Cauchy-Schwarz; die zweite ist die Vertauschung des Integrals mit linearen Abbildungen ($x \mapsto \langle x, v \rangle$ ist linear!). Die Ungleichung folgt aus Cauchy-Schwarz. □

Wir sind weniger an Funktionen $X \to \mathbb{R}^n$ als vielmehr an Funktionen $X \to \mathbb{C}$ interessiert. Das geht nun leicht von der Hand: Man identifiziert nämlich $\mathbb{R}^2$ mit $\mathbb{C}$ mittels $\eta : (x, y) \mapsto x + iy$. Das Standardskalarprodukt auf $\mathbb{R}^2$ wird unter diesem Isomorphismus zu $(z, w) \mapsto \Re(z\bar{w})$, und die ℓ^2-Norm zum Absolutbetrag. Aus dem obigen Satz folgt nun sofort, dass das Integral $\mathscr{L}^1(X;\mathbb{C}) \to \mathbb{C}$ die Abschätzung $|\int f| \leq \int |f|$ erfüllt und $\mathbb{R}$-linear ist.

Das Integral ist auch $\mathbb{C}$-linear, also

$$\int_X zf d\mu = z\int_X f d\mu$$

für $f \in \mathscr{L}^1(X;\mathbb{C})$ und $z \in \mathbb{C}$. Dies beweist man mit folgendem Trick: sind $a, b \in \mathbb{R}$, so kommutiert das Diagram

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "\eta"] \arrow[d, "\begin{pmatrix} a & -b \\ b & a \end{pmatrix}"'] & \mathbb{C} \arrow[d, "(a+ib)\cdot"] \\ \mathbb{R}^2 \arrow[r, "\eta"] & \mathbb{C}. \end{tikzcd}$$

Dann benutze Vertauschung des Integrals mit linearen Abbildungen.

Von nun an werden wir mit dem Symbol $\mathscr{L}^1(X)$ stets den Vektorraum der *komplexwertigen* integrierbaren Funktionen bezeichnen; $\mathbb{R}$-wertige Funktionen sind darin enthalten, und die meisten Argumente übertragen sich auf vektorwertige Funktionen.

Die Nullmengen spielen für das Integral eine spezielle Rolle.

Definition 2.23. *Es sei $(X, \mathcal{B}, \mu)$ ein Maßraum und P eine Eigenschaft von Punkten in X. Wir sagen, dass P* fast überall *gilt, wenn $\{x \in X | \neg P(x)\}$ eine Nullmenge ist.*

Definition 2.24. *Eine messbare Funktion $f : X \to [0, \infty]$ oder $f : X \to \mathbb{R}^n$ heißt* Nullfunktion*, falls $f = 0$ fast überall gilt, also, falls $\{x|f(x) \neq 0\}$ eine Nullmenge ist. Mit dem Symbol $\mathscr{N}(X;\mathbb{R}^n)$ bezeichnen wir die Menge aller Nullfunktionen.*

Ist f eine Nullfunktion, so auch $|f|$, und es gilt dann $\int_X |f(x)|d\mu(x) = \int_X f(x)d\mu(x) = 0$. Die Umkehrung gilt ebenfalls, und dafür notieren wir eine einfache, aber nützlich Ungleichung.

Proposition 2.25. *Es sei X ein Maßraum und $f : X \to [0, \infty]$ messbar. Für $0 < c < \infty$ setzen wir $X_c := \{x \in X|f(x) \geq c\}$. Dann gilt, für alle $c > 0$:*

$$\mu(X_c) \leq \frac{1}{c}\int_X f(x)d\mu(x).$$

Beweis. Der Beweis ist nahezu banal. Es gilt nämlich, nach Definition:

$$c\chi_{X_c} \leq \chi_{X_c} f \leq f$$

und daher wegen der Monotonie des Integrals

$$c\mu(X_c) = \int_X c\chi_{X_c} \leq \int_X f.$$

Division durch c zeigt das gewünschte. □

Korollar 2.26. *Es sei $f \in \mathscr{L}^1(X)$ mit $\int_X |f(x)|d\mu(x) = 0$. Dann ist f eine Nullfunktion.*

Beweis. Sei $X_{1/n}$ wie in Proposition 2.25 definiert. Es gilt dann $\mu(X_{1/n}) \leq 0$, also ist $\{x|f(x) \neq 0\} = \cup_{n=1}^{\infty} X_{1/n}$ eine Nullmenge. □

Nun gilt: falls $f \in \mathscr{L}^1(X)$ und falls g eine Nullfunktion ist, so ist $f+g \in \mathscr{L}^1(X)$, und es gilt

$$\int_X f = \int_X f + g; \int_X |f| = \int_X |f + g|.$$

Dies zeigt, dass wir den Wert einer Funktion auf einer Nullmenge nach Belieben abändern können, ohne dass das Integral sich ändert. Noch einen Schritt weitergehend, können wir auch Funktionen betrachten, die überhaupt nur auf $X - S$ definiert sind (S ist eine Nullmenge) und ihnen dennoch ein Integral $\int_X f$ zuweisen.

Korollar 2.27. *Sei $f : X \to [0, \infty]$ mit $\int_X f < \infty$. Dann gilt $f(x) < \infty$ für fast alle $x \in X$.*

Beweis. Es gilt $f^{-1}(\infty) = \cap_{n\geq 1}\{x|f(x) \geq n\}$. Weil $\mu(\{x|f(x) \geq n\}) \leq \frac{1}{n}\int_X f$ nach 2.25, ist $f^{-1}(\infty)$ eine Nullmenge. □

2.4. **Der Satz von der dominierten Konvergenz.**

Satz 2.28 (Dominierte Konvergenz)**.** *Sei $0 \leq g \in \mathscr{L}^1(X;\mathbb{R})$ und $f_n \in \mathscr{L}^1(X;\mathbb{C})$ eine Folge messbarer Funktionen mit $|f_n(x)| \leq g(x)$. Es existiere der punktweise Grenzwert $f(x) := \lim_n f_n(x)$. Dann*

(1) *f ist integrierbar,*
(2) *$\lim_n \int_X |f_n(x) - f(x)|d\mu = 0$,*
(3) *$\lim_n \int_X f_n d\mu = \int_X f d\mu$.*

Beweis. Weil f messbar und $|f(x)| \leq g(x)$ ist f integrierbar. Ferner ist $|f_n(x) - f(x)| \leq |f_n(x)| + |f(x)| \leq 2g(x)$. Weil $\int_X g d\mu < \infty$, gilt nach Fatou's Lemma

$$\int_X 2g(x)d\mu = \int_X \liminf_n (2g(x) - |f_n(x) - f(x)|)d\mu \leq \liminf_n \int_X (2g(x) - |f_n(x) - f(x)|)d\mu =$$

$$= \liminf_n (\int_X 2g(x)d\mu - \int_X |f_n(x) - f(x)|d\mu) = \int_X 2g(x)d\mu + \liminf_n (-\int_X |f_n(x) - f(x)|d\mu) =$$

$$= \int_X 2g(x)d\mu - \limsup_n \int_X |f_n(x) - f(x)|d\mu.$$

Alle Größen sind nichtnegativ und endlich. Also folgt

$$0 \leq -\limsup_n \int_X |f_n(x) - f(x)|d\mu \Rightarrow \limsup_n \int_X |f_n(x) - f(x)|d\mu \leq 0.$$

Es gilt aber ohnehin $0 \leq \int_X |f_n(x) - f(x)|d\mu$ für alle n, und daher

$$0 \leq \liminf_n \int_X |f_n(x) - f(x)|d\mu \leq \limsup_n \int_X |f_n(x) - f(x)|d\mu \leq 0,$$

also $\lim_n \int_X |f_n(x) - f(x)|d\mu = 0$.

Zuletzt gilt

$$|\int_X f_n d\mu - \int_X f d\mu| \leq \int_X |f_n - f|d\mu \to 0.$$

□

Eine Funktion g, wie im Satz über dominierte Konvergenz gefordert, heißt auch *Majorante* für die Folge f_n. Der Satz über dominierte Konvergenz gehört zu den wichtigsten Sätzen der gesamten Analysis. Die gesamte Entwicklung des Lebesgueintegrals ist so gestaltet, dass der Satz stimmt.

Der Satz von der dominierten Konvergenz gehört mit sofortiger Wirkung zum Standardwerkzeugkasten. Als erste Anwendung zeigen wir, dass das aus der Analysis I bekannte Riemann-Integral mit dem Lebesgue-Integral übereinstimmt.

Satz 2.29. *Seien $a < b \in \mathbb{R}$ und sei $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar. Dann ist f Lebesgue-integrierbar, und das Lebesgue-Integral ist gleich dem Riemann-Integral. Genauer gesagt, für $a < b$ gilt:*

$$L - \int_{[a,b]} f(x)d\mu(x) = R - \int_a^b f(x)dx.$$

Hierbei ist μ selbstverständlich das Lebesguemaß auf $\mathbb{R}$.

Die Notation $L - \int$, $R - \int$ wird nur in diesem und in einem folgenden ähnlichen Beweis benutzt. Wir erinnern daran, dass in der Analysis I das Riemann-Integral so definiert worden ist, dass $\int_a^b = -\int_b^a$. Dies war eine reine Konventionssache, und der Zweck dieser Festlegung ist, dass nur so der Hauptsatz der Differential- und Integralrechnung stimmt. Das so für $a > b$ erklärte Riemann-Integral bemerkt die Richtung des Intervalles. Wir treffen dieselbe Vereinbarung für das Lebesgue-Integral, *falls wir die Notation $\int_a^b f(x)d\mu(x)$ benutzen.*

Für den Beweis von Satz 2.29 führen wir eine neue Vokabel ein.

Definition 2.30. *Eine* naive Treppenfunktion $f : \mathbb{R}^d \to \mathbb{C}$ *ist eine Summe $f = \sum_{i=1}^r a_i \chi_{W_i}$, wobei $a_i \in \mathbb{C}$ und $W_i \subset \mathbb{R}^d$ ein abgeschlossener Würfel ist.*

Eine naive Treppenfunktion f auf $\mathbb{R}$ ist dasselbe wie eine Treppenfunktion im Sinne der Analysis I.

Beweis. Zunächst überprüft man ohne Mühe, dass das Lebesgue-Integral von naiven Treppenfunktionen gleich dem Riemann-Integral naiver Treppenfunktionen ist (weil das Maß eines Quaders gleich dem Elementarvolumen ist!). Genauer gesagt gilt

$$L - \int_{[a,b]} g(x)d\mu(x) = R - \int_a^b g(x)dx$$

für naive Treppenfunktionen g. Wenn f Riemann-integrierbar ist, so ist zunächst f beschränkt, sagen wir $|f(t)| \leq C$ für alle $t \in [a,b]$. Des weiteren gibt es naive Treppenfunktionen $-C \leq g_n \leq g_{n+1} \leq \ldots \leq f \leq \ldots \leq h_{n+1} \leq h_n \leq C$ mit $R - \int_a^b (h_n(x) - g_n(x))dx \to 0$ (das war gerade die Definition der Riemann-Integrierbarkeit). Die Definition des Riemann-Integrals lautet nun

$$R - \int_a^b f(x)dx := \lim_n R - \int_a^b g_n(x)dx = \lim_n L - \int_{[a,b]} g_n(x)d\mu(x).$$

Die Grenzfunktionen $g = \lim_n g_n$ und $h = \lim_n h_n$ existieren und sind messbar, und es gilt $g \leq f \leq h$. Aus dominierter Konvergenz schließen wir

$$\lim_n L - \int_{[a,b]} g_n(x)d\mu(x) = L - \int_{[a,b]} g(x)d\mu(x).$$

Es bleibt zu zeigen, dass f messbar ist, und dass $L - \int_{[a,b]} g(x)d\mu(x) = L - \int_{[a,b]} f(x)d\mu(x)$ gilt. Aber es gilt

$$L - \int_{[a,b]} h(x) - g(x)d\mu(x) = \lim_n L - \int_{[a,b]} h_n(x) - g_n(x)d\mu(x) = 0$$

nach dominierter Konvergenz (der Integrand ist beschränkt durch die konstante Funktion $2C$, und diese ist auf $[a,b]$ integrierbar, weil das Maß von $[a,b]$ endlich ist), somit ist die nichtnegative Funktion $h-g$ eine Nullfunktion (wegen Proposition 2.25). Wegen der Vollständigkeit des Lebesgue-Maßes ist $f-g$ auch eine Nullfunktion, somit messbar. Also ist f messbar und Lebesgue-integrierbar, weil $|f| \leq C\chi_{[a,b]}$, und es gilt

$$L - \int_{[a,b]} g(x)d\mu(x) = L - \int_{[a,b]} f(x)d\mu(x). \qquad \square$$

2.5. **Arbeiten mit den Konvergenzsätzen: Die Gamma-Funktion.** Wir wollen nun illustrieren, wie man die Konvergenzsätze benutzen kann, anhand von parameterabhängigen uneigentlichen Riemann-Integralen. Wir nehmen uns außerdem ein konkretes Beispiel vor, die *Gamma*-Funktion.

Definition 2.31. *Sei $s \in (0,\infty)$. Wir definieren die* Gammafunktion *als*

$$\Gamma(s) := \int_0^\infty t^{s-1}e^{-t}dt.$$

Das Integral ist hierbei als uneigentliches Riemann-Integral zu verstehen, also als

$$\lim_{x\downarrow 0}\int_x^1 t^{s-1}e^{-t}dt + \lim_{y\uparrow\infty}\int_1^y t^{s-1}e^{-t}dt.$$

Ziel ist es, zu zeigen, dass dieses uneigentliche Riemann-Integral tatsächlich ein Lebesgue-Integral ist, und dann mit den Konvergenzsätzen zu zeigen, dass $s \mapsto \Gamma(s)$ eine beliebig oft stetig differenzierbare Funktion ist. Zunächst zeigen wir, dass die Grenzwerte beide existieren. Dafür schätzen wir ab

$$\int_x^1 t^{s-1}e^{-t}dt \leq \int_x^1 t^{s-1}dt = \frac{1}{s}(1-x^s) \to \frac{1}{s}$$

sowie ($m \in \mathbb{N}$, $m \geq s-1$)

$$\int_1^y t^{s-1}e^{-t}dt \leq \int_1^m t^m e^{-t}dt \leq \int_0^\infty t^m e^{-t}dt = m!$$

Das letze Integral ist in Analysis I berechnet worden. Aus diesem Grund sind beide Integrale beschränkt, und weil der Integrand positiv ist, folgt aus dem Satz, dass monotone beschränkte Folgen konvergent sind, dass die Grenzwerte

$$\lim_{x\downarrow 0}\int_x^1 t^{s-1}e^{-t}dt, \lim_{y\uparrow\infty}\int_1^y t^{s-1}e^{-t}dt$$

(in $\mathbb{R}$!) existieren. Daher ist $\Gamma(s)$ eine wohlbestimmte Zahl in $\mathbb{R}$. Die Bedeutung der Γ-Funktion ergibt sich aus folgender Rechnung. Es ist

$$s\Gamma(s) = \int_0^\infty st^{s-1}e^{-t}dt = \int_0^\infty \frac{d}{dt}(t^s)e^{-t}dt.$$

Solche Integrale berechnet man mit partieller Integration. Das Ergebnis ist

$$\int_0^\infty \frac{d}{dt}(t^s)e^{-t}dt = \lim_{x\downarrow 0}\int_x^1 \frac{d}{dt}(t^s)e^{-t}dt + \lim_{y\uparrow\infty}\int_1^y \frac{d}{dt}(t^s)e^{-t}dt =$$

$$= -\lim_{x\downarrow 0}\int_x^1 t^s\frac{d}{dt}(e^{-t})dt - \lim_{y\uparrow\infty}\int_1^y t^s\frac{d}{dt}(e^{-t})dt + \lim_{x\downarrow 0}[t^s e^{-t}]_x^1 + \lim_{y\uparrow\infty}[t^s e^{-t}]_1^y =$$

$$= \Gamma(s+1) + \lim_{x\downarrow 0} x^s e^{-x} + \lim_{y\uparrow\infty} y^s e^{-y} = \Gamma(s+1).$$

Beide Grenzwerte sind Null, der erste weil $s > 0$ und der zweite, weil die Exponentialfunktion stärker fällt als jede Potenz. Wir haben also die *Funktionalgleichung der Gammafunktion*

$$\Gamma(s+1) = s\Gamma(s)$$

bewiesen. Zusammen mit der leichten Rechnung $\Gamma(1) = \int e^{-t}dt = 1$ sehen wir durch Induktion, dass für alle $n \in \mathbb{N}$ gilt

$$\Gamma(n) = (n-1)!$$

Also interpoliert die Gammafunktion die Fakultät, und aus diesem Grund ist die Gammafunktion in manchen Kontexten wichtig. Wir werden sie später im Zusammenhang mit der Berechnung des Maßes der Einheitsbälle in $\mathbb{R}^n$ sehen.

Unser Ziel ist es, zu zeigen, dass die Gammafunktion $\Gamma : (0, \infty) \to \mathbb{R}$ beliebig oft stetig differenzierbar ist. Der erste Schritt besteht darin, das uneigentliche Riemann-Integral als Lebesgue-Integral aufzufassen.

Lemma 2.32. *Seien $a < c < b \in [-\infty, \infty]$ und sei $f : (a, b) \to \mathbb{R}$ stetig. Angenommen, die Grenzwerte*

$$\lim_{x\downarrow a} \int_x^c |f(t)|dt \text{ und } \lim_{y\uparrow b} \int_c^y |f(t)|dt$$

existieren (in $\mathbb{R}$). dann ist $f : (a, b) \to \mathbb{R}$ Lebsegue-integrierbar, und es gilt

$$L-\int_a^b f(t)d\mu(t) = \lim_{x\downarrow a} \int_x^c f(t)dt + \lim_{y\uparrow b} \int_c^y f(t)dt.$$

Beweis. Weil f stetig ist, ist f messbar. Wähle Folgen $x_n \leq c \leq y_n$ mit $x_n \downarrow a$, $y_n \uparrow b$. Dann gilt jedenfalls $\chi_{[x_n,y_n]}(t)f(t) \to f(t)$ punktweise. Außerdem ist $\chi_{[x_n,y_n]}(t)|f(t)| \uparrow |f(t)|$, also liegt monotone Konvergenz vor. Mit dem Satz von Beppo-Levi sehen wir

$$L-\int_a^b |f(t)|d\mu(t) = \lim_{n\to\infty} L-\int_a^b \chi_{[x_n,y_n]}(t)|f(t)|d\mu(t) = \lim_{n\to\infty} L-\int_{x_n}^{y_n} |f(t)|d\mu(t) \overset{2.29}{=}$$

$$= \lim_{n\to\infty} R-\int_{x_n}^{y_n} |f(t)|dt = \lim_{n\to\infty} R-\int_{x_n}^{c} |f(t)|dt + \lim_{n\to\infty} R-\int_{c}^{y_n} |f(t)|dt.$$

Die Existenz dieser Grenzwerte ist vorausgesetzt. Somit ist $L-\int_a^b |f(t)|d\mu(t) < \infty$, also ist f Lebesgue-integrierbar. Ferner gilt $\chi_{[x_n,y_n]}(t)|f(t)| \leq |f(t)|$, also ist $|f(t)|$ eine integrierbare Majorante für die punktweise konvergente Funktionenfolge $\chi_{[x_n,y_n]}f$. Aus dem Satz über dominierte Konvergenz folgt also

$$L-\int_a^b f(t)d\mu(t) = \lim_{n\to\infty} L-\int_a^b \chi_{[x_n,y_n]}(t)f(t)d\mu(t) = \lim_{n\to\infty} L-\int_{x_n}^{y_n} f(t)d\mu(t) \overset{2.29}{=}$$

$$= \lim_{n\to\infty} R-\int_{x_n}^{y_n} f(t)dt = \lim_{n\to\infty} R-\int_{x_n}^{c} f(t)dt + \lim_{n\to\infty} R-\int_{c}^{y_n} f(t)dt,$$

und das ist gerade das uneigentliche Riemannintegral, nach Definition. □

Hieraus folgt, mit den obigen Rechnungen, sofort, dass $f(t) = t^{s-1}e^{-t}$ integrierbar ist und dass $\Gamma(s) = L-\int_0^\infty t^{s-1}e^{-t}dt$. Ab jetzt spielt die Unterscheidung von Riemann- und Lebesgueintegral keine Rolle mehr.

Nun zur Stetigkeit. Sei $s \in (0,\infty)$ und $s_n \to s$ eine konvergente Folge. Wir wollen $\Gamma(s_n) \to \Gamma(s)$ zeigen, und dürfen voraussetzen, dass $0 < \frac{s}{2} \leq s_n \leq 2s < \infty$ für alle n gilt. Es gilt

$$\lim_n t^{s_n-1}e^{-t} = t^{s-1}e^{-t}$$

für alle t, denn die Funktion $s \mapsto t^{s-1} = e^{(s-1)\log(t)}$ ist stetig. Wenn wir den Satz über dominierte Konvergenz anwenden könnten, würden wir

$$\lim_{n\to\infty} \Gamma(s_n) = \lim_{n\to\infty} \int_0^\infty t^{s_n-1}e^{-t}dt = \int_0^\infty \lim_{n\to\infty} t^{s_n-1}e^{-t}dt = \int_0^\infty t^{s-1}e^{-t}dt = \Gamma(s)$$

erhalten, und hätten die Stetigkeit der Γ-Funktion bewiesen. Um die Anwendung des Satzes über dominierte Konvergenz zu rechtfertigen, benötigen wir eine integrierbare Majorante, also eine Funktion $h : (0,\infty) \to \mathbb{R}$, so dass

$$\forall n,t : |t^{s_n-1}e^{-t}| \leq h(t)$$

und

$$\int_0^\infty h(t)dt < \infty$$

gilt. Wir setzen

$$h(t) := \begin{cases} t^{2s-1}e^{-t} & t \geq 1 \\ t^{\frac{s}{2}-1} & t < 1. \end{cases}$$

Nach den oben durchgeführten Rechnungen ist h integrierbar. Für $t < 1$ gilt

$$0 \leq t^{s_n-1}e^{-t} \leq t^{s_n-1} \leq t^{\frac{s}{2}-1} = h(t)$$

und für $t \geq 1$ gilt

$$0 \leq t^{s_n-1}e^{-t} \leq t^{2s-1}e^{-t} = h(t).$$

Somit ist h eine integrierbare Majorante für die Funktionenfolge $t^{s_n-1}e^{-t}$, und die Anwendung des Satzes über dominierte Konvergenz ist gerechtfertigt.

Dieses Argument hat eine abstrakte Version.

Satz 2.33. *Sei X ein Maßraum und Y ein metrischer Raum. Sei $f : X \times Y \to \mathbb{R}^n$ eine Funktion. Es gelte:*

1. *Für jedes $x \in X$ ist die Funktion $f_x : Y \to \mathbb{R}^n$, $f_x(y) := f(x,y)$ stetig.*
2. *Für jedes $y \in Y$ ist die Funktion $f^y : X \to \mathbb{R}^n$, $f^y(x) = f(x,y)$ integrierbar.*
3. *Zu jedem $y \in Y$ gibt es eine Umgebung $U \subset Y$ und eine integrierbare Funktion $g : X \to [0,\infty]$, so dass für alle $(x,z) \in X \times U$ gilt: $|f(x,z)| \leq g(x)$.*

Dann ist die Funktion $F : Y \to \mathbb{R}^n$, welche durch $F(y) := \int_X f(x,y)dx$ definiert ist, stetig.

Beweis. Sei $y_n \to y$ eine konvergente Folge. Wir müssen $\lim_{n\to\infty} F(y_n) = F(y)$ zeigen. Für große n gilt $y_n \in U$, und dann gilt

$$|f(x,y_n)| \leq g(x),$$

und damit mit dominierter Konvergenz:

$$\lim_{n\to\infty} F(y_n) = \lim_{n\to\infty} \int_X f(x,y_n)dx = \int_X \lim_{n\to\infty} f(x,y_n)dx = \int_X f(x,y)dx = F(y).$$

Die vorletzte Gleichung gilt, weil f_x stetig ist. $\square$

Nun zur Differenzierbarkeit. Hierbei ist es hilfreich, das Konvergenzargument zunächst abstrakt zu betrachten.

Satz 2.34 (Ableiten unter dem Integralzeichen)**.** *Es sei X ein Maßraum und $U \subset \mathbb{R}$ offen sowie $f : X \times U \to \mathbb{R}^n$ eine Funktion. Es gelte:*

(1) *Für jedes $x \in X$ ist die Funktion $t \mapsto f(x,t)$ stetig differenzierbar.*
(2) *Für jedes $t \in U$ sind die Funktionen $x \mapsto f(x,t)$ und $x \mapsto \frac{\partial}{\partial t} f(x,t)$ auf X integrierbar.*
(3) *Für jedes $t \in U$ gibt es eine integrierbare Funktion $g : X \to [0,\infty]$ und $\delta > 0$, so dass $(t-\delta, t+\delta) \subset U$ und für $|s-t| < \delta$ sowie alle $x \in X$ gilt: $|\frac{\partial}{\partial t} f(x,t)| \leq g(x)$.*

Dann ist die durch $F(t) := \int_X f(x,t)dx$ definierte Funktion $U \to \mathbb{R}^n$ differenzierbar, und es gilt

$$F'(t) = \int_X \frac{\partial}{\partial t} f(x,t)dx.$$

Beweis. Wir dürfen $n = 1$ annehmen und berechnen mit dem Mittelwertsatz

$$\frac{1}{h}(F(t+h) - F(t)) = \int_X \frac{1}{h}(f(x,t+h) - f(x,t))dx = \int_X \frac{\partial}{\partial t} f(x, t + \theta_x h)dx,$$

wobei $\theta_x \in (0,1)$ von x, t und h abhängen mag. Die Voraussetzung besagt, dass für genügend kleines h gilt: $|\frac{\partial}{\partial t} f(x, t + \theta_x h)| \leq g(x)$. Ferner gilt für jedes x: $\lim_{h\to 0} \frac{\partial}{\partial t} f(x, t+\theta_x h) = \frac{\partial}{\partial t} f(x,t)$, denn die Ableitung ist stetig. Wir dürfen dominierte Konvergenz anwenden und erhalten das gewünschte Ergebnis. $\square$

Ohne neue Ideen lässt sich der Satz auf den Fall $U \subset \mathbb{R}^d$ und partielle Ableitungen verallgemeinern. Wir wenden nun Satz 2.34 auf die Gammafunktion an. Man rechnet

$$\frac{d}{ds} t^{s-1} e^{-t} = \frac{d}{ds} e^{(s-1)\log(t)} e^{-t} = \log(t) t^{s-1} e^{-t}.$$

Sei nun $u \in (0,\infty)$ fest. Für $u/2 \leq s \leq 2u$ und $t \geq 1$ gilt dann

$$|\log(t) t^{s-1} e^{-t}| \leq t^s e^{-t} \leq t^{2u} e^{-t}$$

(beachte, dass $|\log(t)| \leq t$ für $t \geq 1$ gilt). Für $t \leq 1$ gilt

$$|\log(t) t^{s-1} e^{-t}| \leq |\log(t)| t^{s-1} \leq |\log(t)| t^{u/2-1}.$$

Weil

$$\int_0^1 |\log(t)| t^{u/2-1} dt \overset{t=e^{-x}}{=} \int_0^\infty x e^{-x(u/2-1)} e^{-x} dx < \infty$$

ist schließlich

$$g(t) := \begin{cases} |\log(t)| t^{u/2-1} & t < 1 \\ t^{2u} e^{-t} & t \geq 1 \end{cases}$$

eine Majorante für $\frac{d}{ds} t^{s-1} e^{-t}$, $s \in (u/2, 2u)$. Also dürfen wir unter dem Integral ableiten, und erhalten außerdem die Formel

$$\Gamma'(s) = \int_0^\infty \log(t) t^{s-1} e^{-t} dt.$$

Der Vorgang lässt sich wiederholen, und wir finden am Ende:

$$(\frac{d}{ds})^k \Gamma(s) = \int_0^\infty \log(t)^k t^{s-1} e^{-t} dt$$

für alle $k \in \mathbb{N}$.

2.6. **Anwendung auf unendliche Reihen.** Die folgenden Sätze hätten auch in der Analysis I bereits vorkommen können/sollen. Die Beweise werden mit Hilfe der Integrationstheorie jedoch durchsichtiger. Wir betrachten $\mathbb{N}_0$ mit dem Zählmaß. Offensichtlich ist jede Funktion $a : \mathbb{N}_0 \to \mathbb{C}$ messbar.

Satz 2.35. *Die Funktion $a : \mathbb{N}_0 \to \mathbb{C}$ ist genau dann integrierbar, wenn die unendliche Reihe $\sum_{n=0}^{\infty} a(n)$ absolut konvergiert. Das Integral $\int \mathbb{N}_0 a(n) d\mu(n)$ ist gleich der Summe $\sum_{n=0}^{\infty} a(n)$.*

Beweis. Sei $\chi_n : \mathbb{N}_0 \to [0, \infty)$ die charakteristische Funktion der Menge $\{0, 1, \dots, n\}$. Aus dem Satz von Beppo-Levi folgt

$$\int_{\mathbb{N}_0} |a(k)| d\mu(k) = \lim_{n\to\infty} \int_{\mathbb{N}_0} \chi_n(k)|a(k)|d\mu(k) = \lim_{n\to\infty} \sum_{k=0}^{n} |a(k)| = \sum_{k=0}^{\infty} |a(k)|,$$

die zweite Gleichung folgt sofort aus der Definition des Zählmaßes. Daraus folgt die Äquivalenz der absoluten Konvergenz und der Integrierbarkeit. Falls $\sum_{k=}^{\infty} |a(k)| < \infty$, so schließen wir mit dem Satz über dominierte Konvergenz

$$\int_{\mathbb{N}_0} f(k) d\mu(k) = \lim_{n\to\infty} \int_{\mathbb{N}_0}^{\infty} \chi_n(k) f(k) d\mu(k) = \lim_{n\to\infty} \sum_{k=0}^{n} a(k). \qquad \square$$

Falls $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ bijektiv ist, so können wir aus der Reihe $\sum_{n=0}^{\infty} a(n)$ die neue Reihe $\sum_{n=0}^{\infty} a(\varphi(n))$ machen. Diese heißt eine *Umordnung* der ursprünglichen Reihe. Es stellt sich die Frage, ob die Konvergenz von $\sum_{n=0}^{\infty} a(n)$ die Konvergenz der umgeordneten Reihe nach sich zieht und ob die Grenzwerte gleich sind. Der berühmte *Riemannsche Umordnungssatz* besagt, dass dies im allgemeinen falsch ist: ist $a(n) \in \mathbb{R}$ und $\sum_{n=0}^{\infty} a(n)$ konvergent, aber nicht absolut konvergent, so existiert *zu jedem $A \in [-\infty, \infty]$ eine bijektive Abbildung $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ mit $\sum_{n=0}^{\infty} a(\varphi(n)) = A$*! Für absolut konvergente Reihen gibt es aber kein Problem.

Satz 2.36. *Es sei $\sum_{n=0}^{\infty} a(n)$ absolut konvergent und $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ bijektiv. Dann ist $\sum_{n=0}^{\infty} a(\varphi(n))$ absolut konvergent und es gilt $\sum_{n=0}^{\infty} a(\varphi(n)) = \sum_{n=0}^{\infty} a(n)$.*

Beweis. Sei $\eta_n : \mathbb{N}_0 \to [0, \infty)$ die charakteristische Funktion von $\{\varphi(0), \dots, \varphi(n)\}$, also $\eta_n = \chi_n \circ \varphi$ in der Notation des Beweises von Satz 2.35. Weil φ surjektiv ist, gilt $\eta_n \to 0$ (monoton). Mit Beppo-Levi folgt dann

$$\sum_{n=0}^{\infty} |a(\varphi(n))| = \lim_{n\to\infty} \sum_{k=0}^{n} \eta_n(k)|a(\varphi(k))| = \lim_{n\to\infty} \sum_{k=0}^{n} \chi_n(k)|a(k)| = \sum_{k=0}^{\infty} |a(k)|;$$

in der zweiten Gleichung wird die Injektivität von φ verwendet. Hieraus folgt die absolute Konvergenz der umgeordneten Reihe. Die Gleichheit der Grenzwerte folgt unter Berufung auf den Satz über dominierte Konvergenz, wenn man in der obigen Rechnung die Betragsstriche weglässt. $\square$

Auch die Behandlung von Produkten von Reihen wird mit den Konvergenzsätzen sehr einfach.

Satz 2.37. *Seien $\sum_{k=0}^{\infty} a(k)$ und $\sum_{l=0}^{\infty} b(l)$ zwei absolut konvergente Reihen. Dann ist die Funktion $\mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{C}$, $(k, l) \mapsto a(k)b(l)$ integrierbar, und es gilt*

$$\int_{\mathbb{N}_0\times\mathbb{N}_0} a(k)b(l) d\mu(k, l) = (\sum_{k=0}^{\infty} a(k))(\sum_{l=0}^{\infty} b(l)).$$

Beweis. Es sei $\chi_n : \mathbb{N}_0 \times \mathbb{N}_0 \to [0, \infty)$ die charakteristische Funktion von $\{0, \dots, n\} \times \{0, \dots, n\}$. Aus Beppo-Levi folgt

$$\int_{\mathbb{N}_0 \times \mathbb{N}_0} a(k)b(l)d\mu(k,l) = \lim_{n\to\infty} \sum_{k=0}^{n} \sum_{l=0}^{n} |a(k)||b(l)| = (\lim_{n\to\infty} \sum_{k=0}^{n} |a(k)|)(\lim_{n\to\infty} \sum_{l=0}^{n} |b(l)|),$$

was die Integrierbarkeit der Produktreihe zeigt. Die Gleichheit der Grenzwerte folgt dann aus dominierter Konvergenz. □

Satz 2.38. *Seien* $\sum_{k=0}^{\infty} a(k)$ *und* $\sum_{l=0}^{\infty} b(l)$ *zwei Reihen. Das* Cauchy-Produkt *der beiden Reihen ist die Reihe* $\sum_{m=0}^{\infty}$, $c(m) := \sum_{k=0}^{m} a(k)b(m-k)$. *Falls* $\sum_{k=0}^{\infty} a(k)$ *und* $\sum_{l=0}^{\infty} b(l)$ *absolut konvergent sind, so auch das Cauchy-Produkt und es gilt*

$$(\sum_{k=0}^{\infty} a(k))(\sum_{l=0}^{\infty} b(l)) = \sum_{m=0}^{\infty} c(m).$$

Beweis. Sei $\eta_n : \mathbb{N}_0 \times \mathbb{N}_0 \to [0, \infty)$ die Funktion

$$\eta_n(k,l) = \begin{cases} 1 & k + l \le n \\ 0 & k + l > n. \end{cases}$$

Aus dominierter Konvergenz und Satz 2.37 folgt

$$(\sum_{k=0}^{\infty} a(k))(\sum_{l=0}^{\infty} b(l)) = \lim_{n\to\infty} \int_{\mathbb{N}_0 \times \mathbb{N}_0} \eta_n(k,l)a(k)b(l).$$

Es gilt aber $\int_{\mathbb{N}_0 \times \mathbb{N}_0} \eta_n(k,l)a(k)b(l) = \sum_{m=0}^{n} a(k)b(l)$. □

2.7. **Vektorräume von integrablen Funktionen.** Sei $(X, \mathcal{B}, \mu)$ ein Maßraum. Wir führen nun eine ganze Familie von Vektorräumen integrierbarer Funktionen auf X ein.

Definition 2.39. *Sei* $p \in [1, \infty)$. *Sei* $(X, \mathcal{B}, \mu)$ *ein Maßraum und* $f : X \to \mathbb{C}$ *messbar. Wir definieren die* L^p-Norm *von* f *als*

$$\|f\|_p := (\int_X |f(x)|^p d\mu(x))^{1/p} \in [0, \infty].$$

Es sei $\mathscr{L}^p(X)$ *die Menge aller messbaren* f *mit* $\|f\|_p < \infty$. *Sei* $\mathscr{N}(X)$ *die Menge aller Nullfunktionen, d.h. aller* f *mit* $f(x) = 0$ *fast überall.*

Satz 2.40. $\mathscr{L}^p(X)$ *ist ein Vektorraum, und* $\|\|_p$ *ist eine* Halbnorm*, das heißt, für* $f, g \in \mathscr{L}^p(X)$ *und* $a \in \mathbb{C}$ *gilt:*

(1) $\|f\|_p \ge 0$.
(2) $\|af\|_p = |a|\|f\|_p$.
(3) $\|f + g\|_p \le \|f\|_p + \|g\|_p$.

Ist $\|f\|_p = 0$*, so gilt* $f \in \mathscr{N}(X)$.

Bemerkung: für $X = \{1, \dots, n\}$ ist $\mathscr{L}^p(X) = \mathbb{R}^n$, und die L^p-Norm ist gleich der ℓ^p-Norm aus Analysis II. Das deutet bereits an, dass der Fall $p = 2$ ein besonderer ist. Es gibt ein Analogon für den Fall $p = \infty$. Hier ist $\mathscr{L}^\infty(X)$ der Vektorraum aller beschränkten messbaren Funktionen, und $\|f\|_\infty := \inf_{g\in\mathscr{N}(X)} \sup_{x\in X} |f(x)+g(x)|$. Wir wollen den Fall $p = \infty$ nicht betrachten. Die Argumente sind in diesem Fall viel einfacher, aber anders. Überhaupt interessieren wir uns eigentlich nur für die Fälle $p = 1, 2$.

Beweis von Satz 2.40, bis auf die Dreiecksungleichung. Korollar 2.26 zeigt sofort, dass $\|f\|_p = 0 \Rightarrow f \in \mathscr{N}(X)$. Es ist klar, dass $\|f\|_p \geq 0$ und dass $\|af\|_p = |a|\|f\|_p$ gilt. Dass $\mathscr{L}^1(X)$ ein Vektorraum ist, haben wir in Satz 2.22 schon gesehen. Für $p = 1$ ist außerdem die Dreiecksungleichung trivial: $\int |f+g| \leq \int |f| + |g|$.

Wir zeigen nun, dass $\mathscr{L}^p(X)$ ein Vektorraum ist, das heißt: $f, g \in \mathscr{L}^p$, dann $f + g \in \mathscr{L}^p$. Für $a, b \geq 0$ gilt

$$(a+b)^p \leq (2\max\{a,b\})^p = 2^p \max\{a^p, b^p\} \leq 2^p(a^p + a^p).$$

Hieraus folgt: sind $f, g \in \mathscr{L}^p(X)$, dann ist

$$\|f+g\|_p^p = \int_X (|f(x) + g(x)|)^p dx \leq \int_X (|f(x)| + |g(x)|)^p dx \leq$$

$$\leq 2^p \int_X (|f(x)|^p + |g(x)|^p) dx = 2^p(\|f\|_p^p + \|g\|_p^p) < \infty.$$

Somit ist $\mathscr{L}^p(X)$ ein Vektorraum. □

Es bleibt die Dreiecksungleichung im Fall $p \in (1, \infty)$ zu zeigen. Wir geben einen separaten Beweis für $p = 2$, welcher viel einfacher und auch viel wichtiger ist. Die Räume $\mathscr{L}^p(X)$ für $p \neq 1, 2$ gehören in den Bereich des Spezialwissens, aber $\mathscr{L}^2(X)$ sollte jeder Student kennen.

Definition 2.41. *Seien $f, g \in \mathscr{L}^2(X)$. Dann ist das L^2-Skalarprodukt $\langle f, g\rangle$ gegeben durch*

$$\langle f, g\rangle := \int_X f(x)\overline{g(x)} d\mu(x) \in \mathbb{C}.$$

Wir machen uns noch klar, warum die Funktion $f\bar{g}$ überhaupt integrierbar ist. Für $z, w \in \mathbb{C}$ gilt jedenfalls

$$|zw| = |z||w| \leq \frac{1}{2}(|z|^2 + |w|^2)$$

(dies folgt aus der simplen Ungleichung $(a-b)^2 \geq 0$ für $a, b \in \mathbb{R}$). Damit ist $\int_X |f(x)g(\bar{x})| d\mu(x) \leq \frac{1}{2}(\|f\|_2^2 + \|g\|_2^2) < \infty$. Offensichtlich ist

$$\|f\|_2 = \sqrt{\langle x, x\rangle}.$$

Das Skalarprodukt hat alle Eigenschaften, welche von einem komplexen Skalarprodukt gefordert werden, bis auf eine. Ohne Schwierigkeiten rechnet man nach, dass gilt

(1) $\mathbb{R}$-bilinear,
(2) $\langle af, bg\rangle = a\bar{b}\langle f, g\rangle$,
(3) $\langle f, f\rangle \geq 0$,
(4) $\langle f, g\rangle = \overline{\langle g, f\rangle}$.

Was fehlt, ist die Definitheit $\langle f, f\rangle \Rightarrow f = 0$. Es gilt nämlich nur $\langle f, f\rangle = 0 \Rightarrow f \in \mathscr{N}(X)$.

Dennoch gilt aber die *Cauchy-Schwarz-Ungleichung*

$$|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2. \tag{2.42}$$

In der Analysis II wurde die Ungleichung nur für reelle Skalare gezeigt. In der Linearen Algebra II für komplexe Skalare. Der Beweis aus dem Buch von Fischer überträgt sich ohne Schwierigkeiten auf den nicht positiv definiten Fall.

Beweis von (2.42). Man wähle $a \in \mathbb{C}$ mit $|a| = 1$ und $\langle f, ag\rangle \geq 0$. Dann gilt, für alle $t \in \mathbb{R}$:

$$0 \leq \langle f+tag, f+tag\rangle = \langle f,f\rangle + t^2\langle ag, ag\rangle + 2t\Re\langle f, ag\rangle = \langle f,f\rangle + t^2\langle g,g\rangle + 2t\langle f, ag\rangle$$

Falls $\langle g,g\rangle = 0$, so folgt $\langle f,f\rangle + 2t\langle f,ag\rangle \geq 0$ für alle $t \in \mathbb{R}$, was nur für $0 = \langle f, ag\rangle = \bar{a}\langle f,g\rangle$ gehen kann. Ansonsten setze $t = -\frac{\langle f,ag\rangle}{\langle g,g\rangle}$ und es folgt

$$0 \leq \langle f,f\rangle + \frac{\langle f,ag\rangle^2}{\langle g,g\rangle^2}\langle g,g\rangle - 2\frac{\langle f,ag\rangle}{\langle g,g\rangle}\langle f,ag\rangle = \langle f,f\rangle - \frac{\langle f,ag\rangle^2}{\langle g,g\rangle}$$

und somit $\langle f,ag\rangle^2 \leq \langle f,f\rangle\langle g,g\rangle$. □

Explizit hingeschrieben, lautet die Cauchy-Schwarz-Ungleichung übrigens:

$$|\int_X f(x)g\bar(x)d\mu| \leq (\int_X |f(x)|^2 d\mu)^{1/2}(\int_X |g(x)|^2 d\mu)^{1/2}$$

und sieht hochgradig nichttrivial aus. In einer konkreten Anwendung zu erkennen, dass man die Cauchy-Schwarz-Ungleichung anwenden kann, ist nicht immer ganz einfach.

Aus der Cauchy-Schwarz-Ungleichung folgt sofort die Dreiecksungleichung, wie aus LA II und Ana II bekannt. Auch diesen Beweis rekapitulieren wir kurz. Es gilt nämlich

$$\|f+g\|_2 = \langle f+g, f+g\rangle = \|f\|_2^2 + \|g\|_2^2 + 2\Re\langle f,g\rangle \leq \|f\|_2^2 + \|g\|_2^2 + 2|\langle f,g\rangle| \leq$$
$$\leq \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2 = (\|f\|_2 + \|g\|_2)^2.$$

Nun zeigen wir die Dreiecksungleichung für den Fall $p \neq 1, 2$. Im Folgenden sei q immer der *konjugierte Exponent*, d.h.

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Zum Beispiel sind $p = q = 2$ konjugierte Exponenten.

Lemma 2.43 (Youngsche Ungleichung)**.** *Sind $a, b \geq 0$, dann $ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$.*

Beweis. Ist $a = 0$ oder $b = 0$, so ist die Aussage trivial. Für $a, b > 0$ nutzt man die Konvexität der Exponentialfunktion: für $x, y \in \mathbb{R}$ und $t \in [0,1]$ gilt $e^{tx+(1-t)y} \leq te^x + (1-t)e^y$. Setze $t = \frac{1}{p}$, dann ist $(1-t) = \frac{1}{q}$. Sei $x = p\log(a)$ und $y = q\log(b)$. Einsetzen zeigt

$$ab = e^{\frac{1}{p}p\log(a) + \frac{1}{q}q\log(b)} \leq \frac{1}{p}e^{p\log(a)} + \frac{1}{q}e^{q\log(b)} = \frac{1}{p}a^p + \frac{1}{q}b^q.$$

□

Satz 2.44 (Hölder-Ungleichung)**.** *Es gilt*

$$\|fg\|_1 \leq \|f\|_p\|g\|_q.$$

Beweis. Falls $\|f\|_p = 0$ oder $\|g\|_q = 0$, so ist die Ungleichung banal. Wenn $c, d > 0$ und wenn die Ungleichung für f, g gilt, dann auch die für die skalierten Funktionen cf und dg; denn:

$$\|cf dg\|_1 = cd\|fg\|_1;\ \|cf\|_p\|dg\|_q = cd\|f\|_p\|g\|_q.$$

Also reicht es, denn Fall $\|f\|_p = \|g\|_q = 1$ zu studieren. In diesem Fall gilt nach Young

$$|f(x)g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q.$$

Integration über X zeigt dann

$$\|fg\|_1 = \int_X |f(x)g(x)| \leq \frac{1}{p}\int_X |f(x)|^p dx + \frac{1}{q}\int_X |g(x)|^q dx = \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = 1 = \|f\|_p \|q\|.$$

□

Satz 2.45 (Minkowski-Ungleichung)**.** *Es gilt*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Beweis. Berechne

$$\|f+g\|_p^p = \int_X |f+g|^p \leq \int_X |f||f+g|^{p-1} + \int_X |g||f+g|^{p-1} = \|f(f+g)^{p-1}\|_1 + \|g(f+g)^{p-1}\|_1 \leq$$

$$\leq \|f\|_p \|(f+g)^{p-1}\|_q + \|g\|_p \|(f+g)^{p-1}\|_q = (\|f\|_p + \|g\|_p)\|(f+g)^{p-1}\|_q.$$

Andererseits ist $1/p + 1/q = 1$, also $1/q = 1 - 1/p = \frac{p-1}{p}$, also $(p-1)q = p$. Also gilt für jede Funktion h $(= f+g)$

$$\|h^{p-1}\|_q = (\int_X |h|^{(p-1)q})^{1/q} = ((\int_X |h|^p)^{1/p})^{p/q} = \|h\|_p^{p/q}.$$

Somit ist

$$\|f+g\|_p^p \leq (\|f\|_p + \|g\|_p)(\|f+g\|_p)^{p/q},$$

also

$$\|f+g\|_p^{p-p/q} \leq \|f\|_p + \|g\|_p.$$

Aber $p - p/q = p(1-1/q) = 1$. Das zeigt die Ungleichung. □

Wir fügen noch eine nützliche Folgerung der Hölder-Ungleichung ein:

Satz 2.46. *Seien $f, g : X \to [0,\infty]$ messbar und $p \in (1,\infty)$. Dann gilt*

$$(\int_X fg d\mu)^p \leq \|f\|_1^{p-1} \int_X fg^p d\mu.$$

Beweis. Wir definieren ein neues Maß ν auf X durch $\nu(S) := \int_S f d\mu$. Es gilt dann mit der Hölder-Ungleichung

$$(\int_X fg d\mu)^p = (\int_X g d\nu)^p = (\int_X g 1 d\nu)^p \leq (\int_X g^p d\nu)(\int_X 1^q d\nu)^{p/q} =$$

$$(\int_X g^p f d\mu)(\int_X f d\mu)^{p/q},$$

was wegen $p/q = p(1-1/p) = p-1$ den Beweis beendet. □

Weil die L^p-Norm auf $\mathscr{L}^p(X)$ nicht definit ist, ist $\mathscr{L}^p(X)$ kein normierter Vektorraum. Dennoch haben wir den Begriff einer L^p-Cauchyfolge: eine Folge $(f_n)_n$, $f_n \in \mathscr{L}^p(X)$ heißt Cauchyfolge, falls für jedes $\epsilon > 0$ ein n_0 existiert, so dass für $m, n \geq n_0$ gilt: $\|f_n - f_m\|_p \leq \epsilon$. Wir können auch sagen, was Konvergenz bedeutet: $\lim_n \|f_n - f\|_p = 0$. Weil die Norm nicht definit ist, sind Grenzwerte jedoch nicht eindeutig: falls $\|f_n - f\|_p \to 0$ und $\|g - f\|_p = 0$, so gilt wegen der Dreiecksungleichung $\|f_n - g\|_p \to 0$, also ist auch g ein Grenzwert der Folge f_n. Mit anderen Worten: jede Funktion g, welche mit der Grenzfunktion f fast überall übereinstimmt, ist ebenfalls ein Grenzwert. Diese formale Schwierigkeit wollen wir später behandeln. Zunächst das Hauptergebnis.

Satz 2.47 (Normkonvergenzsatz, Satz von Fischer-Riesz)**.** *Es sei X ein Maßraum und $p \in [1, \infty)$.*

(1) *Sei $f_n \in \mathscr{L}^p(X)$ eine L^p-Cauchyfolge und $g \in \mathscr{L}^p(X)$. Falls f_n punktweise fast überall gegen g konvergiert, so gilt $\lim_n \|f_n - g\|_p = 0$.*

(2) *Sei $f_n \in \mathscr{L}^p(X)$ eine L^p-Cauchyfolge. Dann gibt es eine Teilfolge f_{n_k}, welche punktweise fast überall gegen eine Funktion $f \in \mathscr{L}^p(X)$ konvergiert. Es gilt dann $\|f_n - f\|_p \to 0$.*

Beweis. (1): Sei $\epsilon > 0$. Nach Konstruktion existiert ein n_0, so dass für alle $m, n \geq n_0$ gilt $\|f_m - f_n\|_p \leq \epsilon$. Eine Anwendung von Fatou's Lemma zeigt dann

$$\|f - f_n\|_p^p = \int_X |f(x) - f_n(x)|^p dx = \int_X \lim_m |f_m(x) - f_n(x)|^p dx \leq$$

$$\leq \liminf_m \int_X |f_m(x) - f_n(x)|^p dx = \liminf_m \|f_m(x) - f_n(x)\|_p^p \leq \epsilon^p.$$

(2): Sei f_n eine L^p-Cauchyfolge. Wir wählen eine Teilfolge f_{n_k} mit

$$\|f_{n_k} - f_{n_l}\|_p \leq \frac{1}{2^k}$$

für $l \geq k$. Sei

$$g_m(x) := \sum_{k=1}^{m-1} |f_{n_{k+1}}(x) - f_{n_k}(x)|; \; g(x) := \lim_m g_m(x) \in [0, \infty].$$

Auf jeden Fall ist $g : X \to [0, \infty]$ eine messbare Funktion. Wegen der Minkowski-Ungleichung gilt $\|g_m\|_p \leq 1$ für jedes m. Also folgt

$$\int_X g(x)^p dx = \int_X \lim_m g_m(x)^p dx \leq \liminf_m \int_X g_m^p(x) dx \leq 1$$

mit Fatou's Lemma. Deshalb ist $g(x)^p < \infty$ fast überall (2.27), und also ist $S = \{x \in X | g(x) = +\infty\}$ eine Nullmenge. Die Reihe

$$f_{n_m}(x) = f_{n_1}(x) + \sum_{k=1}^{m-1} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

konvergiert punktweise fast überall, nämlich auf $X \setminus S$, gegen eine Funktion f, welche auf der Ausnahmemenge S durch 0 fortgesetzt werde. Weil $|f(x)| \leq g(x)$, ist $f \in \mathscr{L}^p(X)$. Nach Teil (1) gilt $\|f_{n_k} - f\|_p \to 0$. Weil die urprüngliche Folge eine Cauchyfolge war, gilt dann auch $\|f - f_n\|_p \to 0$. Dies sieht man mit einem Argument ein, dass bereits vertraut ist: Beweis des Cauchy-Konvergenzkriteriums (Analysis I) und Beweis, dass folgenkompakte metrische Räume vollständig sind (Analysis II). □

Der Normkonvergenzsatz ist nach dem Satz über dominierte Konvergenz der zweite zentrale Satz der allgemeinen Integrationstheorie. Übrigens ist in Satz 2.47 der Übergang zu einer Teilfolge keineswegs redundant, wie man mit folgendem Beispiel einsehen kann. Sei $X = [0, 1]$, mit dem Lebensgue-Maß. Wir definieren eine Folge f_n mittels $(n \geq 0, k = 0, \ldots, 2^n - 1)$

$$f_{2^n + k} = \chi_{[\frac{k}{2^n}, \frac{k+1}{2^n}]}.$$

Es gilt dann $\|f_{2^n+k}\|_1 = 2^{-n}$. Der Grenzwert in L^1 ist die 0. Aber $f_k(x)$ konvergiert an *keiner* Stelle gegen 0 (an manchen konvergiert es gegen 1). Man mache sich an diesem Beispiel klar, was die Wahl der Teilfolge, die im Beweis getroffen wurde, bewirkt.

Die nächsten beiden Sätze verschaffen uns einen Überblick darüber, welche Funktionen in $\mathscr{L}^p(X)$ liegen.

Satz 2.48 (1. Approximationssatz)**.** *Sei X ein Maßraum und $p \in [1,\infty)$. Sei $f \in \mathscr{L}^p(X)$. Dann gibt es eine L^p-Cauchyfolge von Treppenfunktionen f_n auf X, welche punktweise fast überall und in der L^p-Norm gegen f konvergiert.*

Beweis. Nach Satz 2.12 gibt es eine Folge f_n von Stufenfunktionen mit $|f_n(x)| \leq |f(x)|$ und $\lim_n f_n(x) = f(x)$ für alle x. Weil $\int_X |f_n(x)|^p \leq \|f\|_p^p < \infty$ ist f_n dann eine Treppenfunktion (und nicht nur eine Stufenfunktion), siehe Lemma 2.25. Ferner ist $f_n \in \mathscr{L}^p(X)$. Nun ist

$$|f_n(x) - f(x)|^p \leq 2^p(|f_n(x)|^p + |f(x)|^p) \leq 2 \cdot 2^p |f(x)|^p$$

und $\int_X 2 \cdot 2^p |f(x)|^p < \infty$. Also folgt aus dominierter Konvergenz

$$0 = \int_X \lim_n |f_n(x) - f(x)|^p = \lim_n \int_X |f_n(x) - f(x)|^p = \lim_n \|f_n - f\|_p^p.$$

Insbesondere ist f_n eine L^p-Cauchyfolge. □

Falls X eine offene Teilmenge des $\mathbb{R}^d$ mit dem Lebesgue-Maß ist, so können wir die Aussage von Satz 2.48 noch verschärfen. Hierzu einige Definitionen:

Definition 2.49. *Sei $U \subset \mathbb{R}^d$. Eine* naive Treppenfunktion *auf U ist eine Funktion der Form $f(x) = \sum_{i=1}^r a_i \chi_{W_i}$, wo $W_i \subset U$ ein kompakter Würfel ist. Sei $f : U \to \mathbb{C}$ eine Funktion. Dann ist der* Träger $\operatorname{supp}(f)$ *von f der Abschluss (in U) der Menge $\{x \in U | f(x) \neq 0\}$. Wir sagen, dass f eine stetige Funktion mit kompaktem Träger ist, falls f stetig ist und* $\operatorname{supp}(f) \subset U$ *kompakt ist.*

Satz 2.50 (2. Approximationssatz)**.** *Sei $U \subset \mathbb{R}^d$ offen und $p \in [1,\infty)$. Sei $f \in \mathscr{L}^p(U)$. Dann gilt:*

(1) *Es gibt eine L^p-Cauchy-Folge von naiven Treppenfunktionen f_n auf U, welche punktweise fast überall und in der L^p-Norm gegen f konvergiert.*
(2) *Es gibt eine L^p-Cauchy-Folge von stetigen Funktionen f_n mit kompaktem Träger in U, welche punktweise fast überall und in der L^p-Norm gegen f konvergiert.*

Wir beginnen den Beweis mit zwei Lemmas.

Lemma 2.51. *Seien $K \subset U \subset \mathbb{R}^d$, U offen und K kompakt. Dann gibt es eine stetige Funktion f mit kompaktem Träger in U, so dass $f|_K \equiv 1$.*

Beweis. Sei $\eta := \operatorname{dist}(K, U^c) > 0$ (Satz 1.29). Die Funktion $d_K : \mathbb{R}^d \to \mathbb{R}$, $d_K(x) := \inf_{z \in K} |x - z|$ ist stetig, mit folgendem Argument. Seien $x, y \in \mathbb{R}^d$ und $\epsilon > 0$. Wähle $z \in K$ mit $|x - z| \leq d_K(x) + \epsilon$. Es folgt

$$d_K(y) - d_K(x) \leq |y - z| - |x - z| + \epsilon \leq |y - x| + |x - z| - |x - z| + \epsilon = |y - x| + \epsilon.$$

Weil $\epsilon > 0$ beliebig war, folgt $d_K(y) - d_K(x) \leq |y - x|$. Aus Symmetriegründen gilt auch $d_K(x) - d_K(y) \leq |y - x|$, also $|d_K(x) - d_K(y)| \leq |x - y|$. Deshalb ist d_K Lipschitzstetig und insbesondere stetig.

Nun wähle eine stetige Funktion $h : \mathbb{R} \to [0,1]$ mit $h(0) = 1$, $h(t) = 0$ für $t \geq \eta/2$ (zum Beispiel kann h stückweise linear sein). Dann setze $f(x) := h(d_K(x))$. Dies hat alle gewünschten Eigenschaften. □

Lemma 2.52. *Sei $U \subset \mathbb{R}^d$, $S \subset U$ Lebesgue-messbar mit endlichem Maß.*

(1) *Es gibt zu jedem $\epsilon > 0$ eine Menge $K \subset U$, welche eine disjunkte Vereinigung von endlich vielen kompakten Würfeln ist, so dass $\|\chi_S - \chi_K\|_p \leq \epsilon$.*
(2) *Es gibt zu jedem $\epsilon > 0$ eine stetige Funktion f mit kompaktem Träger in U, so dass $\|f - \chi_S\|_p \leq \epsilon$.*

Beweis. (1): Wegen der äußeren Regularität des Lebesgue-Maßes gibt es eine offene Teilmenge $S \subset V \subset U$ mit $\mu(V - S) \leq \epsilon^p/2$. Wegen Satz 1.31 finden wir eine endliche Vereinigung kompakter Würfel $K \subset V$ mit $\mu(V - K) \leq \epsilon^p/2$. Es gilt dann

$$\|\chi_S - \chi_K\|_p \leq \|\chi_S - \chi_V\|_p + \|\chi_V - \chi_K\|_p = (\int_U |\chi_{V \setminus S}|^p)^{1/p} + (\int_U |\chi_{V \setminus K}|^p)^{1/p} =$$

$$= \mu(V - S)^{1/p} + \mu(V - K)^{1/p} \leq \epsilon.$$

(2): Wegen der Regularität des Lebesgue-Maßes existieren Mengen $K \subset S \subset V \subset U$, V offen, K kompakt mit $\mu(V - K) \leq \epsilon^p$. Wegen Lemma 2.51 existiert eine stetige Funktion f mit kompaktem Träger so dass $\chi_K \leq f \leq \chi_V$. Wir sehen nun, weil $\chi_K \leq \chi_S \leq \chi_V$ und $\chi_K \leq f \leq \chi_V$:

$$f - \chi_S \leq \chi_V - \chi_K; \ \chi_S - f \leq \chi_V - \chi_K.$$

Also

$$\|f - \chi_S\|_p^p = \int_X |f - \chi_S|^p \leq \int_X |\chi_V - \chi_K|^p = \mu(V - K) \leq \epsilon^p.$$ □

Beweis von Satz 2.50. (1): Zu jedem $n \in \mathbb{N}$ wähle eine Treppenfunktion g_n mit $\|f - g_n\|_p \leq \frac{1}{2^{n+1}}$, nach Satz 2.48. Nach Lemma 2.52 (1) gibt es dann eine naive Treppenfunktion f_n mit $\|f_n - g_n\|_p \leq \frac{1}{2^{n+1}}$. Mit der Dreiecksungleichung folgt, dass $\|f - f_n\| \leq \frac{1}{2^n}$. Also konvergiert die Folge f_n in der L^p-Norm gegen f und ist insbesondere eine L^p-Cauchy-Folge. Nach dem Normkonvergenzsatz gibt es eine Teilfolge von f_n, welche überdies punktweise fast überall gegen eine Funktion $g \in \mathscr{L}^p$ konvergiert. Es folgt, aus dem Normkonvergenzsatz, dass $\|f - g\|_p$, also $f(x) = g(x)$ fast überall.

(2): Analog, unter Benutzung von Lemma 2.52 (2). □

Wir wollen nun noch den Makel beheben, dass die L^p-Norm keine Norm ist.

Definition 2.53. *Es sei X ein Maßraum und $p \in [1, \infty)$. Der Vektorraum $L^p(X)$ ist der Quotientenraum*

$$L^p(X) := \mathscr{L}^p(X)/\mathscr{N}(X).$$

Mit anderen Worten: Elemente von $L^p(X)$ sind keine Funktionen, sondern *Äquivalenzklassen* von Funktionen, wobei zwei Funktionen übereinstimmen, wenn $f(x) = g(x)$ fast überall gilt. Wir bezeichnen für den Moment die Äquivalenzklasse von $f \in \mathscr{L}^p(X)$ mit $[f] \in L^p(X)$. Es gilt $[f] = [g]$ genau dann, wenn $f(x) = g(x)$ fast überall. Falls $f(x) = g(x)$ fast überall, so gilt

$$\|f\|_p \leq \|f - g\|_p + \|g\|_p = \|g\|_p \leq \|g - f\|_p + \|f\|_p.$$

Aus diesem Grund ist $\|[f]\|_p := \|f\|_p$ wohldefiniert, also unabhängig von der Wahl von $f \in [f]$. Es gilt dann:

Satz 2.54. *Der Vektorraum $L^p(X)$ ist mit der Norm $\|[f]\|_p := \|f\|_p$ ein normierter Vektorraum. Es gilt sogar: $L^p(X)$ ist ein vollständiger normierter Raum.*

Beweis. Wir müssen noch zeigen, dass die Norm definit ist, und dass Cauchy-Folgen konvergieren. Falls $\|[f]\|_p = 0$, so ist $f(x) = 0$ fast überall, also $[f] = 0 \in L^p(X)$. Falls $[f_n]$ eine Cauchyfolge ist, so konvergiert f_n in L^p-Norm und fast überall gegen eine Funktion $f \in \mathscr{L}^p(X)$. Es gilt dann $\|[f_n] - [f]\|_p \to 0$. $\square$

Warnung !! 2.55. *Man kann die Elemente in $L^p(X)$* nicht *ohne weiteres als Funktionen auffassen. Insbesondere ist es sinnlos, für $f \in L^p(X)$ von $f(x)$ zu sprechen!! Wohldefiniert ist nur das Integral $\int_X f(x)$ im Fall $p = 1$.*

Ebenfalls sei davor gewarnt, dass Elemente in $L^p(X)$ nicht integrierbar sein müssen (außer im Fall $p = 1$). Im Fall $p = 2$ ist dann nur das Skalarprodukt $\langle f, g\rangle$ wohldefiniert.

Es folgen noch zwei Vokabeln.

Definition 2.56. *Ein vollständiger normierter Vektorraum V heißt* Banachraum*. Ist darüber hinaus die Norm auf V durch ein Skalarprodukt induziert, so heißt V* Hilbertraum*.*

3. Mehrdimensionale Integralrechnung

Wir haben bislang die Integrationstheorie weitgehend abstrakt entwickelt, für allgemeine Maßräume. Wir haben auch gelernt, dass das eindimensionale Riemann-Integral in den Formalismus des Lebesgue-Integrals passt. Die *Berechnung* von Integralen haben wir stark vernachlässigt. Auch mit der Berechnung von Lebesguemaßen sind wir nicht wirklich weitergekommen. Dies wird nun nachgeholt, und wir wollen Rechentechniken entwickeln, um Integrale von Funktionen auf $\mathbb{R}^n$ auch konkret auszurechnen. Die zwei wichtigen Methoden sind folgende: mit dem *Satz von Fubini* werden Integrale auf $\mathbb{R}^n$ auf iterierte Integrale auf $\mathbb{R}$ zurückgeführt, und für deren Berechnungen haben wir den Hauptsatz aus der Analysis I. Die zweite Methode ist eine Verallgemeinerung der Substitutionsformel in höhere Dimensionen, der *Transformationssatz*.

3.1. **Der Satz von Fubini.** Wir wollen nun mit μ_n das Lebesgue-Maß auf $\mathbb{R}^n$ bezeichnen. Wie rechnet man Lebesgue-Integrale auf $\mathbb{R}^n$ aus? Es ist naheliegend, für eine Funktion f auf $\mathbb{R}^2$ die Formel

$$\int_{\mathbb{R}^2} f(x,y)d\mu_2(x,y) = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)d\mu_1(x))d\mu_1(y)$$

zu vermuten. Ein Physiker oder Ingenieur mag diese Gleichung geradezu als *Definition* des Integrals ansehen. Damit liegt er *fast* richtig, das heißt unter geeigneten Voraussetzungen an f ist die Gleichung richtig. Es gibt zwei Versionen dieses Satzes, eine für nichtnegative Funktionen, und eine für komplexwertige (oder vektorwertige) Funktionen.

Satz 3.1 (Tonelli)**.** *Es sei $f : \mathbb{R}^{d_1+d_2} \to [0,\infty]$ messbar. Dann gilt*

(1) *Für fast jedes $x \in \mathbb{R}^{d_1}$ ist die Funktion $f_x : \mathbb{R}^{d_2} \to [0,\infty]$, $f_x(y) := f(x,y)$, messbar.*
(2) *Die fast überall definierte Funktion $g : \mathbb{R}^{d_1} \to [0,\infty]$, $g(x) := \int_{\mathbb{R}^{d_2}} f_x(y)dy$ ist messbar.*
(3) $\int_{\mathbb{R}^{d_1}} g(x)dx = \int_{\mathbb{R}^{d_1+d_2}} f(x,y)d(x,y)$.

Satz 3.2 (Fubini)**.** *Es sei $f : \mathbb{R}^{d_1+d_2} \to \mathbb{C}$ integrabel. Dann gilt*

(1) *Für fast jedes $x \in \mathbb{R}^{d_1}$ ist die Funktion $f_x : \mathbb{R}^{d_2} \to \mathbb{C}$, $f_x(y) := f(x,y)$, integrierbar.*
(2) *Die fast überall definierte Funktion $g : \mathbb{R}^{d_1} \to \mathbb{C}$, $g(x) := \int_{\mathbb{R}^{d_2}} f_x(y)dy$ ist integrierbar.*
(3) $\int_{\mathbb{R}^{d_1}} g(x)dx = \int_{\mathbb{R}^{d_1+d_2}} f(x,y)d(x,y)$.

Die wahre Kraft entfalten beide Sätze, wenn sie kombiniert angewendet werden.

Korollar 3.3. *Es sei $f : \mathbb{R}^{d_1+d_2} \to \mathbb{C}$ eine messbare Funktion. Man nehme an:*

(1) *Für fast alle $x \in \mathbb{R}^{d_1}$ ist die Funktion $|f_x| : \mathbb{R}^{d_2} \to \mathbb{C}$ integrierbar,*
(2) *die für fast alle x definierte Funktion $g : \mathbb{R}^{d_1} \to [0,\infty]$; $g(x) := \int_{\mathbb{R}^{d_2}} |f(x,y)|dy$ ist integrierbar.*

Dann ist f integrierbar, und es gilt

$$\int_{\mathbb{R}^{d_1+d_2}} f(x,y)d\mu_{d_1+d_2}(x,y) = \int_{\mathbb{R}^{d_1}} (\int_{\mathbb{R}^{d_2}} f(x,y)d\mu_{d_2}(y))d\mu_{d_1}(x).$$

Beweis. Aus dem Satz von Tonelli folgt $\int_{\mathbb{R}^{d_1+d_2}} |f(x,y)| = \int_{\mathbb{R}^{d_1}} (\int_{\mathbb{R}^{d_2}} |f(x,y)|dy)dx < \infty$, und damit ist f integrierbar. Die Formel für die Integrale folgt aus dem Satz von Fubini. □

Man beachte außerdem die Ähnlichkeit des Satzes von Tonelli mit dem Doppelreihensatz 1.7. In der Tat kann man den Doppelreihensatz als Satz von Tonelli für den Maßraum $I \times J$ mit dem Zählmaß auffassen. Und es gibt auch eine Version des Satzes von Fubini für $I \times J$.

Satz 3.4 (Fubini für das Zählmaß)**.** *Es seien I, J Mengen und $a : I \times J \to \mathbb{C}$ sei integrierbar, also $\sum_{(i,j)\in I\times J} |a_{ij}| < \infty$. Dann gilt:*

1. *Für jedes $i \in I$ gilt $\sum_{j\in J} |a_{ij}| < \infty$.*
2. *Setze $b_i := \sum_{j\in J} a_{ij} \in \mathbb{C}$. Dann gilt $\sum_i |b_i| < \infty$.*
3. *Es gilt $\sum_{i\in I} b_i = \sum_{(i,j)\in I\times J} a_{ij}$.*

Eine Vorsicht: die Messbarkeit von f ist eine *Voraussetzung*, keine Schlussfolgerung des Satzes von Tonelli.

Bevor uns dem Beweis der Sätze 3.1 und 3.2 zuwenden, bringen wir einige Anwendungen. Ein Spezialfall des Satzes von Tonelli ist, wenn $f = \chi_S$, $S \subset \mathbb{R}^{m+n}$ messbar. Dieser Satz hat einen eigenen Namen: das *Cavalieri-Prinzip*.

Satz 3.5 (Cavalieri-Prinzip)**.** *Sei $S \subset \mathbb{R}^{m+n}$ Lebesgue-messbar. Sei, für $x \in \mathbb{R}^m$, $S_x := \{y \in \mathbb{R}^n | (x,y) \in S\} \subset \mathbb{R}^n$. Dann ist für fast alle $x \in \mathbb{R}^m$ die Menge $S_x \subset \mathbb{R}^n$ Lebesgue-messbar, und es gilt*

$$\mu_{m+n}(S) := \int_{\mathbb{R}^m} \mu_n(S_x) d\mu_m(x).$$

Analog kann man die Rolle von x und y vertauschen: $S^y := \{x \in \mathbb{R}^n | (x,y) \in S\} \subset \mathbb{R}^n$. Dann ist

$$\mu_{m+n}(S) := \int_{\mathbb{R}^n} \mu_m(S^y) d\mu_n(y).$$

Einfache Beispiele für das Cavalieri-Prinzip ist ein Quader, und wir erhalten das Ergebnis, dass das Lebesgue-Maß eines Quaders gleich dem Elementarvolumen ist. Von großen Interesse ist die Berechnung von $\mu(D^n)$, des Maßes der Einheitskugel. Man kann dies induktiv mit dem Satz von Fubini berechnen, aber wir geben später eine elegantere Herleitung.

Sei $f : \mathbb{R}^n \to [0,\infty)$ messbar. Wir betrachten die Menge

$$A_f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} | 0 < t < f(x)\}.$$

Um zu sehen, dass A_f Lebesgue-messbar ist, betrachte die Hilfsfunktionen $g, h : \mathbb{R}^{n+1} \to \mathbb{R}$, $g(x,t) = t$ und $h(x,t) = t - f(x)$, welche beide messbar sind. Dann ist

$$A_f = g^{-1}(0,\infty) \cap h^{-1}(0,\infty)$$

messbar. Für $x \in \mathbb{R}^n$ ist

$$(A_f)_x := (0, f(x)) \subset \mathbb{R},$$

also

$$\mu_{n+1}(A_f) = \int_{\mathbb{R}^n} \mu_1(A_f)_x d\mu_n x = \int_{\mathbb{R}^n} \mu_1(0, f(x)) d\mu_n x = \int_{\mathbb{R}^n} f(x) d\mu_n x!!$$

Dies *beweist* die anschauliche Erklärung des Integrals als Flächeninhalt. Wir können auch in anderer Richtung rechnen, und sehen $(A_f)^t = \{x|f(x) > t\}$ und damit

$$\mu_{n+1}(A_f) = \int_0^\infty \mu(\{x|f(x) > t\})dt.$$

Im übrigen kann man auch $B_f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R}|0 \le t \le f(x)\}$ betrachten, und erhält $\mu(B_f) = \mu(A_f)$. Insbesondere ist $B_f - A_f$ eine Nullmenge, und damit ist der Graph von f eine Nullmenge.

Ist $f : \mathbb{R}^{n+m} \to \mathbb{C}$ integrierbar, so folgt aus dem Satz von Fubini

$$\int_{\mathbb{R}^n} (\int_{\mathbb{R}^m} f(x,y)dy)dx = \int_{\mathbb{R}^m} (\int_{\mathbb{R}^n} f(x,y)dx)dy.$$

Diese Formel ist die Grundlage für einen Rechentrick, der überraschend mächtig ist. Als Beispiel geben wir eine Anwendung auf die Berechnung uneigentlicher Integrale.

Beispiel 3.6. *Es gilt*

$$\int_0^\infty \frac{1}{x}\sin(x)dx := \lim_{R\to\infty} \int_0^R \frac{1}{x}\sin(x)dx = \frac{\pi}{2}.$$

In diesem Fall ist das Integral übrigens *nicht* absolut konvergent, was man mit der Periodizizät des Sinus und der Divergenz der harmonischen Reihe einsehen kann. Insbesondere ist $\frac{1}{x}\sin(x)$ nicht Lebesgue-integrierbar. Aus der Regel von L'Hopital folgt aber $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$ Wie kann man das obige Integral mit einem Doppelintegral in Verbindung setzen?? Der Schlüssel ist die durch Substitution schnell gezeigte Formel

$$\frac{1}{x} = \int_0^\infty e^{-xy}dy.$$

Daraus folgt

$$\int_0^R \frac{1}{x}\sin(x)dx = \int_0^R (\int_0^\infty e^{-xy}\sin(x)dy)dx.$$

Um den Satz von Fubini anwenden zu können, muss gezeigt werden, dass die Funktion $f(x,y) = e^{-xy}\sin(x)$ auf $[0,R] \times [0,\infty)$ Lebesgue-integrierbar ist. Hier hilft aber der Satz von Tonelli, denn

$$\int_0^R (\int_0^\infty |e^{-xy}\sin(x)|dy)dx = \int_0^R \frac{1}{x}|\sin(x)|dx$$

und das ist auf jeden Fall endlich, denn das Intervall $[0,R]$ ist kompakt. Also rechnet man

$$\int_0^R (\int_0^\infty e^{-xy}\sin(x)dy)dx = \int_0^\infty (\int_0^R e^{-xy}\sin(x)dx)dy.$$

Um weiterzumachen, braucht man eine Stammfunktion von $g(x) := e^{-xy}\sin(x)$. Dies geht mit der Moivre-Formel $\sin(x) = \frac{1}{2i}(e^{ix} - e^{-ix})$. Daraus lässt sich eine Stammfunktion berechnen. Das Ergebnis ist

$$\frac{e^{-xy}}{y^2+1}(-y\sin(x) - \cos(x)).$$

Also

$$\int_0^\infty (\int_0^R e^{-xy}\sin(x)dx)dy = \int_0^\infty [\frac{e^{-xy}}{y^2+1}(-y\sin(x) - \cos(x))]_{x=0}^{x=R}dy =$$

$$= \int_0^\infty (\frac{e^{-Ry}}{y^2+1}(-y\sin(R) - \cos(R)) + \frac{1}{y^2+1})dy.$$

Wegen des bekannten Integrals $\int_0^\infty \frac{1}{y^2+1}dy = \frac{\pi}{2}$ ist dieser Ausdruck gleich

$$= \pi/2 + \int_0^\infty (\frac{e^{-Ry}}{y^2+1}(-y\sin(R) - \cos(R))dy.$$

Die oben behauptete Formel folgt, wenn wir zeigen können, dass $\lim_{R\to\infty} \int_0^\infty (\frac{e^{-Ry}}{y^2+1}(-y\sin(R) - \cos(R))dy = 0$ gilt. Dies machen wir mit dem Satz über dominierte Konvergenz. Zunächst ist

$$\lim_{R\to\infty} \frac{e^{-Ry}}{y^2+1}(-y\sin(R) - \cos(R)) = 0$$

für jedes $y > 0$. Für $y = 0$ existiert der Grenzwert nicht. Um eine Majorante zu finden, rechne

$$|\frac{e^{-Ry}}{y^2+1}(-y\sin(R) - \cos(R)| \leq |\frac{e^{-Ry}}{y^2+1}(y+1)|$$

und dies ist für alle $R > 0$ integrierbar ("Die Exponentialfunktion gewinnt"). Da wir an $R \to \infty$ interessiert sind, dürfen wir $R \geq 1$ annehmen. Daher ist die Majorante integrierbar, und die Anwendung des Satzes über dominierte Konvergenz gerechtfertigt.

3.2. **Der Beweis des Satzes von Fubini.** Der Beweis der Sätze von Fubini und Tonelli wird in einem Aufwasch geführt.

Lemma 3.7. *Satz 3.1 impliziert Satz 3.2.*

Beweis. Sei $f : \mathbb{R}^{m+n} \to \mathbb{C}$ integrierbar. Dann gilt

$$f = (\Re(f))_+ - (\Re(f))_- + i(\Im(f))_+ - i(\Im(f))_i.$$

Alle vier Funktionen sind nichtnegativ und man kann den Satz von Tonelli auf jede dieser Funktionen anwenden (benutze Linearität des Integrals). □

Es sei $\mathcal{T}$ die Menge aller messbaren $f : \mathbb{R}^{m+n} \to [0, \infty]$, welche die Schlussfolgerung von Satz 3.1 erfüllen. Wir müssen zeigen, dass jede messbare Funktion f in $\mathcal{T}$ liegt. Um überhaupt anfangen zu können, muss man wissen, dass $\mathcal{T}$ nicht leer ist.

Lemma 3.8. *Sei $W \subset \mathbb{R}^{m+n}$ ein kompakter Würfel. Dann gilt $\chi_W \in \mathcal{T}$.*

Dies sieht man durch eine direkte Rechnung.

Lemma 3.9. *Seien $f, g \in \mathcal{T}$ und $c \geq 0$. Dann gilt $f + cg \in \mathcal{T}$.*

Das folgt sofort aus der Additivität des Integrals nichtnegativer Funktionen. Der Kern des Beweis ist, dass $\mathcal{T}$ abgeschlossen unter Grenzübergängen ist.

Lemma 3.10.

(1) *Seien $0 \leq f_1 \leq f_2 \leq \dots$ Funktionen in $\mathcal{T}$. Dann ist die Grenzfunktion $f(x) := \lim_n f_n(x)$ in $\mathcal{T}$.*
(2) *Seien $R, C \geq 0$. Seien f_n Funktionen in $\mathcal{T}$, $0 \leq f_n \leq C\chi_{[-R,R]^n}$ und $f_n(x) \to f(x)$ punktweise. Dann gilt $f \in \mathcal{T}$.*

Beweis. 1. Sei $S_n \subset \mathbb{R}^{d_1}$ eine Nullmenge, so dass $(f_n)_x : \mathbb{R}^{d_2} \to [0,\infty]$, $(f_n)_x(y) = f_n(x,y)$ für alle $x \notin S$ messbar ist und sei $S = \cup_n S_n$. Das ist wieder eine Nullmenge. Für jedes $x \notin S$ ist dann $f_x = \lim (f_n)_x$ messbar. Sei $g_n(x) := \int_{\mathbb{R}^{d_2}} (f_n)_x(y)dy$ und $g(x) := \int_{\mathbb{R}^{d_2}} f_x(y)dy$. Dann sind g und g_n auf dem Komplement von S definiert. Nach Beppo-Levi gilt dann $g(x) = \lim_n g_n(x)$, also ist g messbar auf $\mathbb{R}^{d_1} - S$. Zweimalige Anwendung von Beppo-Levi zeigt dann

$$\int_{\mathbb{R}^{d_1+d_2}} f(x,y)d(x,y) = \lim_n \int_{\mathbb{R}^{d_1+d_2}} f_n(x,y)d(x,y) \overset{f_n\in\mathcal{T}}{=} \lim_n \int_{\mathbb{R}^{d_1}} g_n(x)dx = \int_{\mathbb{R}^{d_1}} g(x)dx.$$

2. Dies geht analog, aber diesmal mit dem Satz über dominierte Konvergenz. Wir müssen überprüfen, dass alle auftretenden Funktionenfolgen integrierbare Majoranten haben. Die Funktionenfolge f_n auf $\mathbb{R}^{d_1+d_2}$ besitzt die integrierbare Majorante $C\chi_{[-R,R]^{d_1+d_2}}$, und es gilt

$$|(f_n)_x| \le C\chi_{[-R,R]^{d_2}}$$

sowie

$$g_n(x) \le C(2R)^{d_2}\chi_{[-R,R]^{d_1}}.$$

□

Leider müssen wir mit den Nullmengen separat zu Rande kommen.

Lemma 3.11. *Sei $S \subset \mathbb{R}^{d_1+d_2}$ eine Nullmenge. Dann gilt $\chi_S \in \mathcal{T}$.*

Beweis. Wir setzen $Y_n := \{x | \mu^*(S_x) \ge \frac{1}{n}\}$. Wir behaupten, dass Y_n eine Nullmenge ist (für jedes n). Es folgt dann, dass für fast alle $x \in \mathbb{R}^{d_1}$ S_x eine Nullmenge ist, und das impliziert die Behauptung auf dem Fuße.

Sei $m \in \mathbb{N}$ beliebig. Wähle abgeschlossene Würfel W_k mit $S \subset \cup_{k=1}^{\infty} W_k$ und $\sum_{k=1}^{\infty} \mu(W_k) \le \frac{1}{mn}$. Setze $f = \sum_{k=1}^{\infty} \chi_{W_k}$. Dann gilt $\chi_S \le f$ und $f \in \mathcal{T}$ nach Lemma 3.8, 3.9 und 3.10. Also gilt

$$\int_{\mathbb{R}_1^d} (\int_{\mathbb{R}^{d_2}} f(x,y)dy)dx = \int_{\mathbb{R}^{d_1+d_2}} f(x,y)d(x,y) \le \frac{1}{mn}.$$

Setze $g(x) = \int_{\mathbb{R}^{d_2}} f(x,y)dy$ (dies ist hier *überall definiert*). Sei $X_n := \{x | g(x) \ge \frac{1}{n}\}$. Es gilt dann $Y_n \subset X_n$. Wegen Proposition 2.25 gilt

$$\frac{1}{n}\mu(X_n) \le \frac{1}{mn} \Rightarrow \mu^*(Y_n) \le \mu(X_n) \le \frac{1}{m}.$$

Weil m beliebig war, sehen wir $\mu^*(Y_n) = 0$, wie behauptet. □

Lemma 3.12. *Jede Nullfunktion $f : \mathbb{R}^{d_1+d_2} \to [0,\infty]$ liegt in $\mathcal{T}$.*

Denn eine solche Nullfunktion ist aufsteigender Limes von Treppenfunktionen, welche alle Nullfunktionen sind. Solche sind in $\mathcal{T}$, nach Lemma 3.11 und 3.9.

Beweis von Satz 3.1. Sei zunächst $f : \mathbb{R}^{d_1+d_2} \to [0,\infty]$ messbar, und $f \le C\chi_{[-R,R]^{d_1+d_2}}$. Dann ist f integrierbar, und nach Satz 2.50 finden wir eine L^1-Cauchyfolge f_n von naiven Treppenfunktionen, welche fast überall und in L^1-Norm gegen f konvergiert. Ferner können wir die f_n so wählen, dass $0 \le f_n \le C\chi_{[-2R,2R]^{d_1+d_2}}$. Es gibt dann Nullfunktionen g_n, g, so dass $f_n - g_n \to f - g$ überall und in Norm.

Die naiven Treppenfunktionen f_n gehören zu $\mathcal{T}$, nach Lemma 3.8, 3.9, und daher auch $f_n - g_n$, nach Lemma 3.8. Aus Lemma 3.10 folgt dann $f - g \in \mathcal{T}$, und weil $g \in \mathcal{T}$, auch $f \in \mathcal{T}$.

Sei $f : \mathbb{R}^{d_1+d_2} \to [0,\infty]$ messbar. Sei

$$f_n := \min(\chi_{[-n,n]^{d_1+d_2}} f, n).$$

Dann ist nach dem ersten Beweisschritt $f_n \in \mathcal{T}$, es gilt $0 \leq f_1 \leq f_2 \leq \ldots \to f$, und aus Lemma 3.10 folgt dann $f \in \mathcal{T}$. □

Es ist nicht immer einfach, die Messbarkeit einer Funktion auf $\mathbb{R}^{m+n}$ zu beweisen. Insbesondere ist folgende Aussage falsch.

Falsche Aussage 3.13. *Sei $S \subset \mathbb{R}^{m+n}$. Angenommen, für fast alle $x \in \mathbb{R}^m$ sei $S_x \subset \mathbb{R}^n$, $S_x = \{y|(x,y) \in S\}$ messbar, und die Funktion $\mathbb{R}^m \to [0,\infty]$, $x \mapsto \mu(S_x)$ sei messbar. Dann ist S messbar.*

Gegenbeispiele 3.14. *Sei $T \subset \mathbb{R}$ nicht messbar. Sei $S_x \subset \mathbb{R}$ durch*

$$S_x := \begin{cases} [0,1) & x \in T \\ [1,2) & x \notin T \end{cases}$$

gegeben. Betrachte $S = \cup_{x\in\mathbb{R}} \{x\} \times S_x \subset \mathbb{R}^2$. Dann erfüllt S die Voraussetzungen der obigen (falschen) Aussage, denn jedes S_x ist messbar mit Maß 1. Dass S nicht messbar ist, sieht man ein, indem man die Schnitte S^y in die andere Richtung bildet ($S^y = \{x|(x,y) \in S\}$. Dann ist $S^y = T$, falls $y \in [0,1)$ und $S^y = T$ für $y \in [1,2)$. Nach dem Satz von Tonelli ist dann S nicht messbar.

Eine positive und machmal nützliche Aussage ist folgede.

Satz 3.15. *Seien $f : \mathbb{R}^m \to \mathbb{C}$ und $g : \mathbb{R}^n \to \mathbb{C}$ messbar. Dann ist die Funktion $h : \mathbb{R}^{m+n} \to \mathbb{C}$, $(x,y) \mapsto f(x)g(y)$ messbar.*

Beweis. Seien $p_1 : \mathbb{R}^{m+n} \to \mathbb{R}^m$ und $p_2 : \mathbb{R}^{m+n} \to \mathbb{R}^n$ die Projektionen. Die Funktion h kann als Produkt $(f \circ p_1)(g \circ p_2)$ geschrieben werden. Also genügt es, zu zeigen, dass p_1 und p_2 $\mathcal{L}$-$\mathcal{L}$-messbar sind. Betrachte $p = p_1$. Nach satz 1.43 genügt es, zu beweisen, dass die Urbilder Von Borelmengen unter p wieder Borelmengen sind (das ist klar, weil p stetig ist) und das die Urbilder von Nullmengen unter p wieder Nullmengen sind. Sei $S \subset \mathbb{R}^m$ eine Nullmenge. Wir behaupten, dass $T := p^{-1}(S) \subset \mathbb{R}^{m+n}$ eine Nullmenge ist. Offenbar ist $T = S \times \mathbb{R}^n$. Wegen der σ-Subadditivität des äußeren Maßes reicht es, zu beweisen, dasss $T_k := S \times [-k,k]^n$ eine Nullmenge ist. Sei $\epsilon > 0$. Überdecke S mit abzählbar vielen Würfeln in $\mathbb{R}^m$, so dass $\sum_{i=1}^\infty \mu(W_i) \leq \epsilon$. Dann ist T_k in der Vereinigung $\cup_{i=1}^\infty W_i \times [-k,k]^n$ enthalten. Es folgt

$$\mu^*(T_k) \leq \sum_{i=1}^\infty \mu(W_i \times [-k,k]^n) = \sum_{i=1}^\infty (2k)^n \mu(W_i) \leq (2k)^n \epsilon.$$

Folglich, weil das für alle $\epsilon > 0$ gilt, ist T_k eine Nullmenge und damit auch $T = \cup_{k=1}^\infty T_k$. □

3.3. Die Transformationsformel.

Satz 3.16 (Transformationsformel)**.** *Es seien $U, V \subset \mathbb{R}^d$ offen und sei $F : U \to V$ ein C^1-Diffeomorphismus. Dann gilt:*

1. *Eine Menge $A \subset U$ ist genau dann Lebesgue-messbar, wenn $F(A) \subset V$ Lebesgue-messbar ist.*

(2) *Sei $f : V \to [0,\infty]$ messbar. Dann gilt*

$$\int_V f(y)dy = \int_U f(F(x))|\det DF(x)|dx \in [0,\infty].$$

(3) *Sei $f : V \to \mathbb{C}$ integrierbar. Dann ist $U \to \mathbb{C}$, $x \mapsto f(F(x))|\det DF(x)|$ integrierbar, und es gilt*

$$\int_V f(y)dy = \int_U f(F(x))|\det DF(x)|dx \in \mathbb{C}.$$

Hierzu einige Bemerkungen: der erste Teil impliziert, dass die Funktion $f \circ F$ messbar ist, wenn f es ist. Damit ist auch $x \mapsto f(F(x))|\det DF(x)|$ messbar.

Der Beweis ist der schwierigste des ganzen Semesters, und verlangt alles ab, was wir entwickelt haben. Deshalb werden wir vorher die Anwendungen und Beispiele diskutieren.

Sei zunächst $S \subset U$ messbar. Dann gilt $\chi_{F(S)} \circ F = \chi_S$ (sic). Durch Einsetzen in die Transformationsformel erhalten wir

$$\mu(F(S)) = \int_V \chi_{F(S)}(y)dy = \int_U \chi_{F(S)}(F(x))|\det(DF(x))|dx =$$

$$= \int_U \chi_S(x)|\det(DF(x))|dx = \int_S |\det DF(x)|dx.$$

Ein interessanter Fall eines Diffeomorphismus ist eine lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^n$, $F(x) := Ax$ mit $A \in \mathrm{GL}_n(\mathbb{R})$. Für eine messbare Teilmenge $S \subset \mathbb{R}^n$ folgt

$$\mu(F(S)) = \int_S |\det DF(x)|dx = \int_S |\det A|dx = |\det(A)|\mu(S).$$

Ist spezieller $S = [0,1]^n$, so ist also $|\det(A)| = \mu(F([0,1]^n))$. Dies stimmt mit der geometrischen Definition der Determinante von A als Volument des von den Vektoren $Ae_1, \dots, Ae_n$ aufgepannten Spates überein. Man könnte dies auch direkt mit dem Satz von Fubini und einer Faktorisierung von A in ein Produkt von Elementarmatrizen zeigen. Ein raffinierter Grenzübergang impliziert dann die Transformationsformel. Wir werden aber einem anderen Weg folgen, bei dem die Determinante eher indirekt ins Spiel kommt.

3.4. **Integration rotationssymmetrischer Funktionen.** Wir wollen nun die Transformationsformel benutzen und ganz bestimmte Funktionen integrieren. Sei $f : (0,\infty) \to \mathbb{C}$ eine Funktion, und man betrachte die Funktion $g : \mathbb{R}^n \setminus 0 \to \mathbb{C}$, $x \mapsto f(\|x\|_2)$. Dann hängt $g(x)$ nur von $\|x\|$ ab (g is rotationssymmetrisch). Wir werden das Integral von g berechnen, indem wir es auf ein eindimensionales Integral zurückführen. Zunächst betrachtet man den Fall $n = 2$.

Sei $F : U = (0,2\pi) \times (0,\infty) \to \mathbb{R}^2$, $F(t,r) := (r\cos(t), r\sin(t))$. Es gilt dann

$$DF(t,r) = \begin{pmatrix} \cos(t) & -r\sin(t) \\ \sin(t) & r\cos(t) \end{pmatrix};\ \det DF(t,r) = r.$$

Das Bild von F ist die offene Teilmenge $V := \mathbb{R}^2 \setminus \{(x,0)|x \geq 0\}$, und die Umkehrabbildung $F^{-1} : V \to U$ ist wieder eine C^1-Funktion. Somit ist F ein Diffeomorphismus. Ferner sieht man sofort, dass $\mathbb{R}^2 \setminus V$ eine Nullmenge ist. Sei nun $f : (0,\infty) \to \mathbb{C}$ messbar (die genauen Voraussetzungen, wann f integrierbar ist, sehen wir gleich, und es kommt uns hier aufs Rechnen an) sowie $g : \mathbb{R}^2 \setminus 0 \to \mathbb{C}$ durch $g(x) := f(\|x\|)$ definiert. Es gilt dann nach der Transformationsformel

$$\int_{\mathbb{R}^2} g(x)dx = \int_V g(x)dx = \int_U g(F(t,r))|\det DF(t,r)|d(t,r) = \int_U f(r)rd(t,r).$$

Nach Fubini schließen wir weiter

$$\int_U f(r)rd(t,r) = \int_0^\infty (\int_0^{2\pi} rf(r)dt)dr = 2\pi \int_0^\infty rf(r)dr.$$

zum Beispiel können wir $f(r) = \chi_{[0,1]}$ betrachten, und erhalten für das Maß des Einheitskreises die bekannte Formel

$$\mu(D^2) = 2\pi \int_0^1 rdr = \pi.$$

Ein berühmtes Beispiel ist

$$f(r) = e^{-r^2}.$$

Man kann zeigen, dass diese Funktion keine elementar ausdrückbare Stammfunktion besitzt. Dennoch lässt sich das uneigentliche Integral $\int_0^\infty f(r)dr$ nun leicht berechnen. Dafür berechnet man

$$\frac{d}{dr}e^{-r^2} = -2re^{-r^2} \Rightarrow 2\pi \int_0^\infty re^{-r^2}dr = 2\pi\frac{1}{2}e^{-r^2}|_\infty^0 = \pi$$

mit dem Hauptsatz. Sei $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x) = e^{-|x|^2} = e^{-\langle x,x\rangle} = e^{-x_1^2}e^{-x_2^2}$. Nach der oben durchgeführten Rechnung und Fubini ist

$$\pi = \int_{\mathbb{R}^2} e^{-x_1^2}e^{-x_2^2}dx_1dx_2 = (\int_{-\infty}^\infty e^{-x^2}dx)^2.$$

Es folgt die überraschende, von Gauss bewiesene Formel

$$\int_{-\infty}^\infty e^{-x^2}dx = \sqrt{\pi}.$$

Weil e^{-x^2} eine gerade Funktion ist, gilt mit der Substitution $x = \frac{1}{\sqrt{2}}y$ auch

$$\frac{1}{2}\sqrt{\pi} = \int_0^\infty e^{-x^2}dx = \frac{1}{\sqrt{2}}\int_0^\infty e^{-\frac{1}{2}y^2}dy.$$

Mit Fubini sehen wir außerdem

$$\int_{\mathbb{R}^n} e^{-\|x\|^2}dx = \pi^{n/2}. \tag{3.17}$$

Wir wollen nun rotationssymmetrische Funktionen auf $\mathbb{R}^n$, $n \geq 3$ integrieren.

Eine Möglichkeit einer geeigneten Koordinatentransformation bieten die sogenannten Kugelkoordinaten, was aber bereits für $n = 3$ zu eher lästigen Rechnungen führt. Eine geschicktere Variante bietet die *stereographische Projektion*

$$h : \mathbb{R}^{n-1} \to S^{n-1},$$

die wir zunächst beschreiben wollen. Anschaulich berühren sich S^{n-1} und $-e_n + \mathbb{R}^{n-1}$ in dem Punkt $-e_n$. Einen Punkt $x \in \mathbb{R}^{n-1}$ bildet man auf den Schnittpunkt der Gerade zwischen e_n und $x - e_n$ mit S^{n-1}. In Formeln:

$$h(x) := (1-t)e_n + t(x - e_n),$$

wobei $t \in (0,1)$ durch die Forderung $\|(1-t)e_n + t(x-e_n)\|_2^2 = 1$ gegeben ist. Diese Bedingung führt auf die Gleichung

$$1 = \|(1-t)e_n + t(x-e_n)\|^2 = \|(1-2t)e_n + tx\|^2 = (1-2t)^2 + t^2\|x\|^2$$

mit der Lösung

$$t = \frac{4}{4+\|x\|^2}.$$

Die stereographische Projektion ist bijektiv auf $S^{n-1}\backslash\{e_n\}$. An den Formeln erkennt man, dass h stetig differenzierbar ist. Die Umkehrabbildung ist durch die Formel

$$k(y) = y + e_n + \frac{1+\langle y, e_n\rangle}{1-\langle y, e_n\rangle}(y-e_n)$$

gegeben, wie man durch Einsetzen bestätigt.

Nun sei

$$H : (0,\infty) \times \mathbb{R}^{n-1} \to \mathbb{R}^n;\ H(r,x) = rh(x).$$

Dann ist H bejektiv auf $\mathbb{R}^n \setminus [0,\infty]e_n$. Das Komplement des Bildes von H ist eine Nullmenge. Die Umkehrabbildung von H ist durch die Formel

$$H^{-1}(y) = (\|y\|, k(\frac{y}{\|y\|}))$$

gegeben. Dies ist eine Zusammensetzung differenzierbarer Funktionen, also wieder differenzierbar. Folglich ist H ein Diffeomorphismus.

Sei nun wieder $f : (0,\infty) \to \mathbb{C}$ und $g(y) := f(\|y\|)$. Wir wollen g integrieren, und auser Transformationsformel folgt (weil $\|H(r,x)\| = r$), zusammen mit Fubini,

$$\int_{\mathbb{R}^n} f(\|y\|)dy = \int_0^\infty \int_{\mathbb{R}^{n-1}} f(r)|\det DH(r,x)|dxdr.$$

Was ist $DH(r,x)$? Es gilt

$$\frac{\partial}{\partial r}H(r,x) = h(x);\ \frac{\partial}{\partial x_i}H(r,x) = r\frac{\partial}{\partial x_i}h(x).$$

Setze $\Delta(x) := |\det(h(x), D_1h(x), \dots, D_{n-1}h(x))|$. Es folgt

$$|\det DH(r,x)| = r^{n-1}\Delta(x).$$

Eingesetzt ergibt sich

$$\int_{\mathbb{R}^n} f(\|y\|)dy = \int_0^\infty \int_{\mathbb{R}^{n-1}} f(r)|\det DH(r,x)|dxdr = \int_0^\infty \int_{\mathbb{R}^{n-1}} f(r)r^{n-1}\Delta(x)dxdr =$$

$$(\int_{\mathbb{R}^{n-1}} \Delta(x)dx)(\int_0^\infty f(r)r^{n-1}dr) =: C_n(\int_0^\infty f(r)r^{n-1}dr).$$

Es bleibt noch die Konstante C_n zu bestimmen. Hierfür muss man die Ableitung von h ausrechnen, in die Determinante einsetzen. Die Formel für h deutet an, dass dies alles andere als ein Vergnügen ist. Selbst wenn die Berechnung gelingt, ist am Ende das Integral über $\mathbb{R}^{n-1}$ zu berechnen. Man kann versuchen, zu nutzen, dass $\Delta(x)$ wieder rotationssymmetrisch ist, aber das sieht alles andere als erfolgsversprechend aus. Es gibt einen *viel* geschickteren Weg, und die Rechnung ist ein Dreizeiler. *Bevor man stur etwas ausrechne, denke man immer darüber nach, welche Information genau nötig ist.*

Wir brauchen nämlich nur eine Funktion $f:(0,\infty)\to\mathbb{R}$, so dass wir $\int_{\mathbb{R}^n} f(\|y\|)dy$ und $\int_0^\infty f(r)r^{n-1}dr$ ausrechnen können (das Integral darf aber nicht Null sein). Ein Versuch ist $f=\chi_{[0,1]}$. Dann finden wir

$$\mu(D^n)=C_n(\int_0^1 r^{n-1}dr)=\frac{C_n}{n},$$

und haben C_n mit dem Volumen der Einheitskugel in Verbindung gebracht, das man mit Fubini ausrechnen kann. Noch geschickter ist es, die Funktion $f(r)=e^{-r^2}$ einzusetzen. Aus 3.17 folgt nämlich

$$\pi^{n/2}=\int_{\mathbb{R}^n} e^{-\|x\|^2}=C_n\int_0^\infty r^{n-1}e^{-r^2}dr.$$

Nun substituiere man $r=\sqrt{t}$ und sieht

$$\pi^{n/2}=C_n\int_0^\infty t^{\frac{1}{2}(n-1)}e^{-t}\frac{1}{2\sqrt{t}}dt=\frac{1}{2}C_n\int_0^\infty t^{\frac{n}{2}-1}e^{-t}dt=:\frac{1}{2}C_n\Gamma(\frac{n}{2}),$$

wobei wir die Definition der Gammafunktion eingesetzt haben. Wir haben somit die Formel

$$C_n=\frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$$

gefunden. Es gilt $\Gamma(\frac{2m}{2})=(m-1)!$, aber können wir auch $\Gamma(n/2)$ berechnen? Die Funktionalgleichung $s\Gamma(s)=\Gamma(s+1)$ der Gammafunktion zeigt

$$\Gamma(\frac{2m+1}{2})=\frac{2m-1}{2}\Gamma(\frac{2m-1}{2})=\ldots=\frac{2m-1}{2}\frac{2m-3}{2}\cdots\frac{1}{2}\Gamma(\frac{1}{2})=\frac{(2m-1)!!}{2^m}\Gamma(\frac{1}{2}),$$

wobei $(2m-1)!!=\prod_{k=0}^{m-1}(2k+1)$ die *Doppelfakultät* ist. Schließlich ist

$$\Gamma(\frac{1}{2})=2\int_0^\infty x^{1-1)}e^{-x^2}dx=\sqrt{\pi}.$$

Wir notieren noch die Formel für $\mu(D^n)$. Es ist

$$\mu(D^n)=\frac{\pi^{n/2}}{\frac{n}{2}\Gamma(\frac{n}{2})}=\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$$

mit der Funktionalgleichung der Gammafunktion. Für $n=2m$ folgt

$$\mu(D^{2m})=\frac{\pi^m}{m!}$$

und für $n=2m-1$

$$\mu(D^{2m-1})=\frac{\pi^{m-1}\sqrt{\pi}}{\Gamma(\frac{2m-1}{2}+1)}=\frac{\pi^{m-1}\sqrt{\pi}}{\Gamma(\frac{2m+1}{2})}=\frac{\pi^{m-1}2^m}{(2m-1)!!}.$$

Der Spezialfall

$$\mu(D^1)=2$$

ist vertrauenerweckend, und die Formel

$$\mu(D^3)=\frac{4\pi}{3}$$

ist aus dem Geometrieunterricht bekannt.

3.5. **Beweis der Transformationsformel.** Nach den hoffentlich etwas unterhaltsamen Berechnungen kommen wir zu der großen Herausforderung, die Transformationsformel zu beweisen. Für den Beweis müssen wir fast alle Register ziehen, die in den letzten 2 1/2 Semestern zusammengekommen sind. Der Beweis, den wir geben, stammt von Bröcker [1], wobei die maßtheoretischen Grundlagen in [1] anders aufgebaut sind, was einige Modifikationen erfordert.

Im wesentlichen basiert der Beweis auf zwei Grundideen. Erstens reduzieren wir den Satz auf einfache Funktionen (charakteristische Funktionen kompakter Würfel), und zweitens führen wir eine Induktion über d durch. Der Fall $d = 1$ ist hierbei ähnlich (aber nicht gleich) wie die Substitutionsregel für das Riemann-Integral. Im Induktionsschritt wird der Umkehrsatz benutzt, um F in eine einfachere Form zu bringen.

Es sei stets $F : U \to V$ ein C^1-Diffeomorphismus offener Teilmengen des $\mathbb{R}^d$. Zuerst beweisen wir die 1. Aussage des Transformationssatzes, nämlich, dass $F(S) \subset V$ genau dann Lebesgue-messbar ist, wenn $S \subset U$ es ist.

Lemma 3.18. *Sei $S \subset U$ eine Nullmenge. Dann ist $F(S) \subset V$ eine Nullmenge.*

Beweis. Wir statten den $\mathbb{R}^d$ mit der Maximumsnorm (bzw. ℓ^∞-Norm) aus. Dann sind Bälle Würfel. Betrachte zunächst einen kompakten Würfel $A \subset U$. Es gibt dann einen kompakten Würfel B mit $A \subset \mathring{B} \subset B \subset U$. Sei $L := \sup_{x \in B} \|DF(x)\| < \infty$. Wir nehmen zunächst an, dass S in A enthalten ist. Sei $\epsilon > 0$. Dann gibt es Würfel W_n, mit $S \subset \cup_{n=1}^\infty W_n$, $W_n \subset B$ und $\sum_{n \geq 1} \mu(W_n) < \epsilon$. Sei a_n die Kantenlänge von W_n, und $x_n \in W_n$ der Mittelpunkt. Dann gilt, für $y \in W_n$: $\|F(z) - F(x_n)\| \leq L\|z - x_n\| \leq La_n$. Also ist $F(W_n)$ in einem Würfel der Kantenlänge La_n enthalten. Daraus folgt

$$\mu(F(W_n)) \leq L^d a_n^d = L^d \mu(W_n).$$

Insgesamt gilt

$$\mu(F(S)) \leq \sum_{n=1}^{\infty} L^d \mu(W_n) \leq \epsilon L^d.$$

Weil $\epsilon > 0$ beliebig war, und L nur von A abhängt, ist $F(S)$ eine Nullmenge, wenn $S \subset A$. Im allgemeinen Fall schreiben wir $U = \cup_{n=1}^\infty W_n$ als fast-disjunkte Vereinigung von abgeschlossenen Würfeln. Es gilt dann

$$\mu(F(S)) \leq \sum_{n=1}^{\infty} \mu(F \cap W_n) = 0.$$

□

Lemma 3.19. *Eine Teilmenge $S \subset U$ ist genau dann messbar, wenn $F(S) \subset V$ messbar ist.*

Beweis. Sei $S \subset U_n \subset U$, U_n offen mit $\mu(U_n - S) \leq 1/n$. Dann ist $T := \cap_n U_n$ eine Borelmenge. Es gilt $S \subset T$ und $T - S$ ist eine Nullmenge. Also ist

$$F(T) = F(S) \cup F(T - S).$$

Weil F ein Diffeomorphismus ist, ist $F(T)$ eine Borelmenge, und nach Lemma 3.18 ist $F(T - S)$ eine Nullmenge. Also ist $F(T)$ Lebesgue-messbar, und somit auch

$$F(S) = F(T) \setminus F(T - S).$$

□

Lemma 3.20. *Angenommen, Satz 3.16 gilt für naive Treppenfunktionen $f : V \to \mathbb{C}$. Dann gilt der Satz für alle integrierbaren Funktionen.*

Beweis. Sei $f : V \to \mathbb{C}$ integrierbar. Dann gibt es, nach Satz 2.50, eine L^1-Cauchyfolge f_n naiver Treppenfunktionen mit Träger in V, welche fast überall auf V gegen f konvergiert. Nach Voraussetzung gilt

$$\int_V |(f_n(y) - f_m(y))| dy = \int_U |f_n(F(x)) - f_m(F(x))|| \det DF(x)| dx.$$

Also ist die Funktionenfolge $g_n(x) := f_n(F(x))| \det DF(x)|$ eine Cauchyfolge in $L^1(U)$. Nach Lemma 3.18 konvergiert g_n dann fast überall gegen $g(x) = f(F(x))| \det DF(x)|$. Aus dem Normkonvergenzsatz folgt, dass $\|g - g_n\|_1 \to 0$. Dies impliziert, dass

$$\int_U f_n(F(x))| \det DF(x)| dx \to \int_U f(F(x))| \det DF(x)| dx.$$

Weil außerdem

$$\int_U f_n(F(x))| \det DF(x)| dx = \int_V f_n(y) dy \to \int_V f(y) dy,$$

folgt, dass Satz 3.16 für integrierbare Funktionen gilt. Für messbare Funktionen $f : V \to [0, \infty]$ argumentiert man wie folgt. Wir wählen nach Satz 1.30 kompakte Würfel $W_j \subset V$, so dass $\cup_{j=1}^{\infty} W_j = V$. Setze $V_k := \cup_{j=1}^{k} W_j$, dies ist eine kompakte Teilmenge von V. Die Funktionen

$$f_k(y) := \chi_{V_k}(y) \min\{k, f(x)\}$$

sind integrierbar, und die Funktionenfolge (f_k) konvergiert monoton gegen f. Weil die Transformationsformel für jede Funktion f_k gilt, können wir mit zweimaliger Anwendung des Satzes von Beppo-Levi auf die Gültigkeit der Transformationsformel für f schließen:

$$\int_V f(y) dy = \lim_{k \to \infty} \int_V f_k(y) dy = \lim_{k \to \infty} \int_U f_k(F(x))| \det DF(x)| dx = \int_U f(F(x))| \det DF(x)| dx.$$

□

Lemma 3.21. *Der Satz 3.16 ist richtig für $d = 1$.*

Beweis. Nach Lemma 3.20 genügt es, dies für naive Treppenfunktionen zu beweisen, und wegen der Linearität des Integrals können wir uns auf eine charakteristische Funktion $f = \chi_{[a,b]}$, $[a, b] \in V$, $a < b$, beschränken. Es gilt aber $\int_V \chi[a, b](y) dy = b - a$. Die Behauptung lautet also

$$b - a = \int_U \chi_{[a,b]}(F(x)| \det DF(x)| dx.$$

Nun ist $\chi_{[a,b]} \circ F = \chi_{[F^{-1}(a), F^{-1}(b)]}$. Wir unterscheiden zwei Fälle. Entweder ist $F^{-1}(a) < F^{-1}(b)$. Dann ist $F|_{[F^{-1}(a), F^{-1}(b)]}$ streng monoton steigend, und hat positive Ableitung. Es gilt dann

$$\int_U \chi_{[a,b]}(F(x))| \det DF(x)| dx = \int_{F^{-1}(a)}^{F^{-1}(b)} F'(x) dx = b - a$$

nach dem Hauptsatz der Integral- und Differentialrechnung. Wenn $F^{-1}(a) > F^{-1}(b)$, dann ist $F|_{[F^{-1}(a), F^{-1}(b)]}$ streng monoton fallend, und hat negative Ableitung. Wir rechnen dann

$$\int_U \chi_{[a,b]}(F(x))|\det DF(x)|dx = -\int_{F^{-1}(b)}^{F^{-1}(a)} F'(x)dx = \int_{F^{-1}(a)}^{F^{-1}(b)} F'(x)dx = b-a.$$

Wir haben hier natürlich auch benutzt, dass die Determinante einer 1×1-Matrix mit Eintrag a gerade a ist. □

Lemma 3.22. *Wenn Satz 3.16 für die Diffeomorphismen $F : U \to V$ und $G : V \to W$ gilt, dann auch für $G\circ F$. Falls F eine Permutation der Koordinaten ist, so gilt Satz 3.16 für F.*

Beweis. Die erste Aussage ist eine direkte Rechnung, benutzend, dass $\det(AB) = \det(A)\det(B)$ für alle Matrizen gilt. Die zweite folgt sofort aus Fubini. □

Allmählich kommen wir nun zum Kern des Beweises. Wir führen einen Induktionsbeweis.

Lemma 3.23. *Angenommen, Satz 3.16 sei für $d-1$ schon gezeigt. Sei $F : U \to V$ ein Diffeomorphismus der Form*

$$F(x_1,\dots,x_d) = (x_1, F_2(x),\dots,F_d(x)).$$

Dann ist der Satz richtig für F.

Beweis. Für $t\in\mathbb{R}$ setze $U_t = \{x\in\mathbb{R}^{d-1} | (t,x)\in U\}$, und V_t analog. Schreibe $F(t,x) = (t, G_t(x))$, wobei $G_t : U_t \to V_t$ ein Diffeomorphismus ist. Es gilt

$$|\det DF(t,x)| = |\det DG_t(x)|.$$

Sei nun $f : V \to \mathbb{C}$ integrierbar. Wir berechnen:

$$\int_V f(t,y)d(t,y) = \int_{\mathbb{R}}(\int_{V_t} f(t,y)dy)dt.$$

Dies folgt aus dem Satz von Fubini. Formal müsste man den Definitionsbereich auf ganz $\mathbb{R}^d$ fortsetzen, durch 0. Die Induktionsannahme impliziert

$$\int_{\mathbb{R}}(\int_{V_t} f(t,y)dy)dt = \int_{\mathbb{R}}(\int_{U_t} f(t,G_t(x))|\det DG_t(x)|dx)dt =$$

$$= \int_{\mathbb{R}}(\int_{U_t} f(t,G_t(x))|\det DF(t,x)|dx)dt = \int_U f(t,G_t(x))|\det DF(t,x)|d(t,x).$$

wobei die letzte Gleichung wieder aus dem Satz von Fubini folgt. Dieses Integral ist aber einfach gleich

$$\int_U f(t,G_t(x))|\det DF(t,x)|d(t,x) = \int_U f(F(t,x))|\det DF(t,x)|d(t,x). \qquad \square$$

Lemma 3.24. *Unter der Annahme, dass Satz 3.16 für $d-1$ bereits gezeigt ist, existieren zu jedem $p\in U$ Umgebungen $U_p\subset U$ und $V_p = F(U_p)\subset V$ von p beziehungsweise $F(p)$, so dass Satz 3.16 für die Einschränkung $F|_{U_p}\to V_p$ gilt.*

Beweis. Betrachte $p\in U$. Nach Komposition mit einer Koordinatenpermutation, welche die Gültigkeit des Transformationssatzes nicht ändert (Lemma 3.22), können wir annehmen, dass

$$\frac{\partial}{\partial x_1}F_1(p)\neq 0$$

gilt. Sei $G : U \to \mathbb{R}^d$ durch
$$G(y) = (F_1(y), y_2, \ldots, y_d)$$
gegeben. Dann gilt
$$\det(DG(y)) = \frac{\partial}{\partial y_1} F_1(y)$$
und damit ist $DG(p)$ invertierbar. Nach dem *Umkehrsatz* aus der Analysis II gibt es eine Umgebung $U_p \subset U$ von p, so dass $G : U_p \to G(U_p) =: W_p$ ein Diffeomorphismus ist. Sei $H : W_p \to V_p := F(U_p)$ die Komposition $F \circ G^{-1}$. Es folgt, dass $F : U_p \to F(U_p)$ als Komposition
$$U_p \stackrel{G}{\to} W_p \stackrel{H}{\to} V_p$$
geschrieben werden kann. Der Diffeomorphismus G hält, nach Anwendung einer Koordinatenpermutation, die erste Koordinate fest. Daher gilt der Transformationssatz für G, wegen Lemma 3.23. Nun seien $H_1, \ldots, H_d$ die Komponenten von H. Berechne $F(y) = H(G(y)) =$
$$= (H_1(F_1(y), y_2, \ldots, y_d), H_2(F_1(y), y_2, \ldots, y_d), \ldots, H_d(F_1(y), y_2, \ldots, y_d)).$$
Hieraus folgt, dass $H_1(F_1(y), y_2, \ldots, y_d) = F_1(y)$ gilt, und das heißt $H(z_1, \ldots, z_d) = (z_1, H_2(z), \ldots, H_d(z))$. Also ist auch H von der speziellen Gestalt. Also gilt die Transformationsformel auch für H. Nach Lemma 3.22 sind wir fertig. □

Jetzt können wir alle Fäden zusammenführen.

Beweis von Satz 3.16, Schluss. Wir führen den Beweis durch Induktion über d. Der Fall $d = 1$ ist in Lemma 3.21 gezeigt worden, und wir nehmen an, der Satz gelte für $d - 1$.

Es genügt, nach Lemma 3.20, den Beweis zu führen, wenn $f = \chi_K$ die charakteristische Funktion eines kompakten Würfels in V ist. Seien $(V_p)_p$ wie in Lemma 3.24 konstruiert. Dies ist eine offene Überdeckung von V und schränkt sich zu einer offenen Überdeckung von K ein. Nach dem Lebesgue-Lemma können wir K in endich viele, disjunkte Teilwürfel $K_1, \ldots, K_r$ einteilen, so dass jeder der Teilwürfel in einer der Mengen V_p zu liegen kommt. Es gilt
$$\chi_K = \chi_{K_1} + \ldots + \chi_{K_r}.$$
Also genügt, es die Funktionen χ_{K_i} zu betrachten. Aber der Transformationssatz gilt für alle diese Funktionen, nach Induktionsannahme und Lemma 3.24. □

4. Faltung und Fouriertransformation

Definition 4.1. *Die* Faltung *zweier Funktionen f, g auf $\mathbb{R}^n$ ist definiert als*

$$f * g(x) := \int_{\mathbb{R}^n} f(x-y)g(y)dy,$$

falls das Integral existiert.

Lemma 4.2. *Sei $f \in L^p(\mathbb{R}^n)$, $p < \infty$. Sei $T_y f(x) := f(x-y)$ die Verschiebung von f. Dann ist die Abbildung $y \mapsto T_y f$ eine stetige Abbildung $\mathbb{R}^n \to L^p(\mathbb{R}^n)$.*

Beweis. Aus der Translationsinvarianz des Lebesgue-Maßes folgt $\|T_y f\|_p = \|f\|_p$. Sei zunächst f stetig mit kompaktem Träger, welcher in $B_R(0)$ enthalten sei. Dann ist f gleichmäßig stetig. Sei $\epsilon > 0$ und sei $\delta > 0$ so gewählt, dass $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$ gilt. Dann gilt für $|y| \le \delta \le R$:

$$\|T_y f - f\|_p^p = \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p dx = \int_{B_{2R}(0)} |f(x-y)-f(x)|^p dx \le \mu(B_{2R}(0))\epsilon^p.$$

Hieraus folgt die Stetigkeit von $y \mapsto T_y f$ bei 0. Für $z \in \mathbb{R}^n$ und $|y| \le \delta$ gilt

$$\|T_{z+y} f - T_z f\| = \|T_z(T_y f - f)\|_p = \|T_y f - f\|_p.$$

Für beliebiges $f \in L^p(\mathbb{R}^n)$ wenden wir den Approximationssatz an. Sei $\epsilon > 0$. Wähle $g \in C_c(\mathbb{R}^n)$ mit $\|f - g\|_p \le \epsilon/3$. Dann gilt

$$\|T_y f - f\|_p \le \|T_y f - T_y g\|_p + \|T_y g - g\|_p + \|g - f\|_p \le \frac{2}{3}\epsilon + \|T_y g - g\|_p.$$

Weil g stetig mit kompaktem Träger ist, konvergiert der letzte Term gegen Null, wie im ersten Beweisteil bewiesen. □

Satz 4.3. *Sei $f \in L^1$, $g \in L^p$. Dann ist die Faltung $f * g$ fast überall definiert, und die Funktion $f * g$ ist L^p, sowie $\|f * g\|_p \le \|f\|_1 \|g\|_p$.*

Beweis. Weil das Produkt einer L^1-Funktion mit einer L^p-Funktion nicht integrierbar zu sein braucht, ist es nicht klar, dass das Integral überhaupt existiert. Auf jeden Fall ist aber die Funktion $h(x,y) := f(x-y)g(y)$ auf $\mathbb{R}^{2n}$ messbar. Hier hilft Satz 3.15, denn die Funktion h, komponiert mit der linearen Funktion $(x,y) \mapsto (x-y, y)$ ist von der Form wie in Satz 3.15. Nun berechne, mit Tonelli,

$$\int (\int |f(x-y)g(y)|dy)^p dx \overset{2.46}{\le} \int\int |f(x-y)||g(y)|^p \|f\|_1^{p-1} dydx =$$

$$\int\int |f(x-y)||g(y)|^p \|f\|_1^{p-1} dxdy = \int |g(y)|^p \|f\|_1 \|f\|_1^{p-1} dxdy = \|g\|_p^p \|f\|_1^p.$$

Dies zeigt, dass $(\int |f(x-y)g(y)|dy)^p < \infty$ für fast alle x, also $\int |f(x-y)g(y)|dy < \infty$ für fast alle x. Ferner ist $f * g \in L^p$, und die gewünschte Normabschätzung ergibt sich ebenso. □

Das Problem, dass die Faltung nicht überall definiert ist, tritt auch wirklich auf. Betrachte etwa $f = \frac{1}{\sqrt{x}}$ und $g = \frac{1}{\sqrt{-x}}$, wobei f auf $[0,1]$ definiert ist, und g auf $[-1,0]$ und beide Funktionen durch Null fortgesetzt werden. Dann sind $f, g \in L^1$, aber nicht in L^2. Dann führt $f * g(0)$ auf das Integral $\int_0^1 \frac{1}{x}dx$.

Einfacher ist die Situation im Fall konjugierter Exponenten.

Satz 4.4. *Sei $f \in L^p$ und $g \in L^q$, wobei wie üblich $\frac{1}{p} + \frac{1}{q} = 1$. Falls $p, q > 1$, so ist die Funktion $f * g$ stetig und beschränkt durch $\|f\|_p \|g\|_q$.*

Beweis. Hier zeigt die Hölder-Ungleichung, dass $\int |f(x-y)g(y)| < \|f\|_p \|g\|_q$, und deshalb ist die Faltung überall definiert. Ferner ist

$$|f * g(x) - f * g(z)| \leq |\int (f(x-y) - f(z-y))g(y)dy| \leq \|T_x f - T_z f\|_p \|g\|_q$$

nach Hölder, und daraus folgt die Stetigkeit nach Lemma 4.2. $\square$

Die Bedeutung der Faltung liegt darin, dass sie eine strukturierte Approximation von L^p-Funktionen erlaubt.

Lemma 4.5. *Es sei* $f : \mathbb{R}^n \to \mathbb{C}$ lokal integrierbar*, d.h.* $\int_{B_R(x)} |f(y)|dy < \infty$ *für jeden Ball in* $\mathbb{R}^n$*. Ferner sei* g *eine beliebig oft stetig differenzierbare Funktion mit kompaktem Träger. Dann ist* $g * f$ *eine beliebig oft stetig differenzierbare Funktion, und es gilt* $\frac{\partial}{\partial x_i}(g * f) = (\frac{\partial}{\partial x_i} g) * f$*.*

Beweis. Das folgt mittels Differentation unter dem Integral. Formal rechnet man

$$\frac{\partial}{\partial x_i}(g * f)(x) = \int_{\mathbb{R}^n} \frac{\partial}{\partial x_i} g(x-y) f(y) dy = (\frac{\partial}{\partial x_i} g) * f(x),$$

und die Differentation unter dem Integral ist zu rechtfertigen. Aber $y \mapsto \frac{\partial}{\partial x_i} g(x-y)$ ist stetig mit kompaktem Träger, und deshalb ist der Integrand $\frac{\partial}{\partial x_i} g(x-y) f(y)$ integrierbar. Falls $\operatorname{supp}(g) \subset B_R(0)$ und $|\frac{\partial}{\partial x_i} g| \leq C$, so ist der Integrand für alle x nahe bei x_0 durch die Funktion

$$\chi_{B_{2R}(x_0)}(y) C f(y)$$

beschränkt, welche nach der Voraussetzung an f gerade integrierbar ist. $\square$

Man kann die Voraussetzungen des Lemmas auf verschiedene Arten variieren, indem man die Voraussetzungen an g abschwächt und die an f verstärkt.

Lemma 4.6. *Es gibt eine beliebig oft stetig differenzierbare Funktion mit kompaktem Träger* $\phi : \mathbb{R}^n \to [0, \infty)$ *so dass* $\int_{\mathbb{R}^n} \phi(x) dx = 1$*.*

Beweis. Man beginnt mit der Funktion $g : \mathbb{R} \to \mathbb{R}$,

$$g(x) := \begin{cases} e^{-\frac{1}{x}} & x > 0 \\ 0 & x \leq 0. \end{cases}$$

Diese Funktion ist beliebig oft stetig differenzierbar. Nun seien $a < b$ und

$$h_{a,b}(x) := \frac{g(x-a)}{g(b-x) + g(x-a)}.$$

Der Nenner ist stets > 0: für $x < b$ ist $g(b-x) > 0$ und für $x > a$ der andere Term. Falls $x < a$, so gilt $h_{a,b}(x) = 0$ und für $x > b$ gilt $h_{a,b}(x) = 1$. Dazwischen ist $0 \leq h_{a,b} \leq 1$. Mit Hilfe der Funktionen $h_{a,b}$ definiere man

$$\tilde{\phi}(x) := 1 - h_{1,2}(\|x\|).$$

Diese Funktion ist C^∞, hat Träger im Ball $\bar{B}_2(0)$. Ferner $\tilde{\phi} \geq 0$ und $\tilde{\phi}(0) = 1$. Insbesondere ist $\int_{\mathbb{R}^n} \tilde{\phi}(x) dx > 0$ und wir setzen

$$\phi(x) := \frac{\tilde{\phi(x)}}{\int_{\mathbb{R}^n} \tilde{\phi}(y) dy}. \qquad \square$$

Solche Hilfsfunktionen sind in vielen Zusammenhängen nützlich. Nun definieren wir, für $t > 0$, $\phi_t(x) := \frac{1}{t^n}\phi(\frac{x}{t})$. Je kleiner t wird, desto mehr konzentiert sich ϕ_t um 0. Das Integral von ϕ_t ist nach der Transformationsformel aber immer gleich 1.

Satz 4.7. *Ist $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, dann gilt*

$$\lim_{t\to 0} \|\phi_t * f - f\|_p = 0.$$

Beweis. Sei zunächst g stetig mit kompaktem Träger, sagen wir $\operatorname{supp}(g) \subset B_{R_1}(0)$. Außerdem sei $\operatorname{supp}(\phi_t) \subset B_R(0)$, für alle $t \le 1$. Dann gilt zumindest $\operatorname{supp}(\phi_t * g) \subset B_{R_1+R}(0)$. Wenn wir zeigen können, dass $\phi_t * g$ gleichmäßig gegen g konvergiert, so ist der Satz für g gezeigt, denn

$$\|\phi_t*g-g\|_p^p = \int_{\mathbb{R}^n} |\phi_t*g(y)-g(y)|^p dy = \int_{B_{R+R_1}(0)} |\phi_t*g(y)-g(y)|^p dy \le \mu(B_{R+R_1}(0))\|\phi_t*g-g\|_\infty.$$

Sei $\epsilon > 0$. Weil g gleichmäßig stetig ist, gibt es ein $\delta > 0$ mit $|x-y| \le \delta \Rightarrow |f(x)-f(y)| \le \epsilon$. Es gilt dann

$$|\phi_t * g(x) - g(x)| = |\int_{\mathbb{R}^n} \phi_t(x-y)g(y)dy - g(x)| = |\int_{\mathbb{R}^n} \phi_t(x-y)(g(y)-g(x))dy| \le$$

$$\le \int_{\mathbb{R}^n} |\phi_t(x-y)(g(y)-g(x))|dy = \int_{\mathbb{R}^n} \phi_t(x-y)|(g(y)-g(x))|dy$$

weil $\int_{\mathbb{R}^n} \phi_t = 1$ und weil $\phi_t \ge 0$. Dieses Integral spalten wir auf als

$$\int_{|x-y|\le\delta} \phi_t(x-y)|(g(y)-g(x))|dy + \int_{|x-y|\ge\delta} \phi_t(x-y)|(g(y)-g(x))|dy.$$

Nun ist der Träger von ϕ_t enthalten im Ball vom Radius $\frac{R}{t}$, also ist das zweite Integral Null, falls $\frac{R}{t} \le \delta$, also für genügend große t, unabhängig von x. Das erste Integral ist aber

$$\int_{|x-y|\le\delta} \phi_t(x-y)|(g(y)-g(x))|dy \le \int_{|x-y|\le\delta} \phi_t(x-y)\epsilon dy \le \epsilon,$$

weil $\int_{\mathbb{R}^n} \phi_t = 1$.

Sei nun $f \in L^p$ beliebig. Sei $\epsilon > 0$. Wähle eine stetige Funktion mit kompaktem Träger g so dass $\|f-g\|_p \le \frac{1}{3}\epsilon$. Es gilt dann

$$\|\phi_t*f-f\|_p \le \|\phi_t*f-\phi_t*g\|+\|\phi_t*g-g\|+\|g-f\| \le \|\phi_t*(f-g)\|+\|\phi_t*g-g\|+\frac{\epsilon}{3},$$

wobei wir die Linearität der Faltung benutzt haben, also die (offensichtliche) Gleichung $\phi_t * g - \phi_t * f = \phi_t * (g-f)$. Es ergibt sich nach Satz 4.3

$$\|\phi_t*(f-g)\|+\|\phi_t*g-g\|+\frac{\epsilon}{3} \le \|\phi_t\|_1\|f-g\|_p+\|\phi_t*g-g\|+\frac{\epsilon}{3} = \|\phi_t*g-g\|+\frac{2\epsilon}{3}.$$

Weil g stetig mit kompaktem Träger ist, konvergiert der erste Term gegen Null, nach dem ersten Teil des Beweises. □

Definition 4.8. *Sei $f \in L^1(\mathbb{R}^n)$. Dann ist die* Fourier-Transformation *von f die Funktion*

$$\hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x)e^{-ix\xi}dx.$$

Dieses Integral ist für alle ξ definiert, weil $|e^{-ix\xi}| = 1$. Die grundlegenden Eigenschaften der Fouriertransformation sind in folgendem Satz zusammengefasst.

Satz 4.9. *Seien $f, g \in L^1$. Dann gilt*

1. *$\hat{f}$ ist stetig und es gilt $|\hat{f}(\xi)| \leq \frac{1}{(2\pi)^{n/2}} \|f\|_1$ für alle ξ.*
2. *Für die verschobene Funktion $T_y f(x) = f(x-y)$ gilt $\widehat{T_y f}(\xi) = e^{-iy\xi} \hat{f}(\xi)$.*
3. *Die Funktion $\hat{f}$ "verschwindet im Unendlichen", das heißt es gilt $\lim_{|\xi| \to \infty} \hat{f}(\xi) = 0$.*
4. *Es gilt $\widehat{f * g}(\xi) = (2\pi)^{n/2} \hat{f}(\xi) \hat{g}(\xi)$.*

Beweis. 1. Der Integrand $x \mapsto f(x) e^{-ix\xi}$ ist stetig in ξ und, unabhängig von ξ, beschränkt durch die integrierbare Funktion $|f(x)|$. Dies sind die Voraussetzungen von Satz 2.33, und die Stetigkeit von $\hat{f}$ folgt. Ferner ist $|\hat{f}| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x) e^{-ix\xi}| dx = \frac{1}{(2\pi)^{n/2}} \|f\|_1$.

2. Einfache Rechnung:

$$\widehat{T_y f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x-y) e^{-ix\xi} dx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x-y) e^{-i(x-y)\xi} e^{-iy\xi} dx = e^{-iy\xi} \hat{f}(\xi).$$

3. Sei $\epsilon > 0$. Wähle $\delta > 0$, so dass $|y| \leq \delta \Rightarrow \|T_y f - f\|_1 \leq \epsilon$, was nach Satz 4.2 geht. Wir nehmen nun die Fouriertransformation dieser Ungleichung. Nach Teil 1 gilt dann für jedes $\xi \in \mathbb{R}^n$:

$$|(e^{-iy\xi} - 1) \hat{f}(\xi)| = |\widehat{T_y f}(\xi) - \hat{f}(\xi)| \leq \frac{1}{(2\pi)^{n/2}} \epsilon$$

für jedes ξ und jedes y mit $|y| \leq \delta$. Sei nun $|\xi| \geq \frac{\pi}{\delta}$ und setze $y := \frac{\pi}{|\xi|^2} \xi$. Dann gilt $|y| \leq \delta$, also

$$\frac{1}{(2\pi)^{n/2}} \epsilon |(e^{-i\pi} - 1) \hat{f}(\xi)| = 2 |\hat{f}(\xi)|,$$

denn $e^{i\pi} = -1$. Daraus folgt die Behauptung.

4. Dies ist eine direkte Rechnung:

$$\widehat{f * g}(\xi) = \frac{1}{(2\pi)^{n/2}} \int f * g(x) e^{-ix\xi} dx = \frac{1}{(2\pi)^{n/2}} \int \int f(x-y) g(y) e^{-ix\xi} dy dx.$$

Weil sowohl f als auch g L^1-Funktionen sind, ist der Integrand in $L^1(\mathbb{R}^n \times \mathbb{R}^n)$ (Tonelli). Wir dürfen also die Integrationsreihenfolge vertauschen und finden

$$\widehat{f * g}(\xi) = \frac{1}{(2\pi)^{n/2}} \int \int f(x-y) g(y) e^{-ix\xi} dx dy = \int \widehat{T_y f}(\xi) g(y) dy = \int \hat{f}(\xi) e^{-iy\xi} g(y) dy = (2\pi)^{n/2} \hat{f}(\xi) \hat{g}(\xi).$$

□

Um die weiteren Eigenschaften der Fouriertransformation zu diskutieren, ist es bequem, einen neuen Funktionenraum einzuführen. Hierfür führen wir die Multiindex-Notation ein. Sei $\alpha = (j_1, \ldots, j_n)$, $j_1, \ldots, j_n \in \mathbb{N}_0$. Dann setze

$$D^\alpha f := D_1^{j_1} \cdots D_n^{j_n} f; \; x^\alpha = x_1^{j_1} cdots x_n^{j_n}.$$

Definition 4.10. *Eine* Schwartz-Funktion $f : \mathbb{R}^n \to \mathbb{C}$ *ist eine Funktion, welche beliebig oft stetig differenzierbar ist und zwar so, dass für alle Multiindices α, β die Funktion*

$$x^\beta D^\alpha f$$

beschränkt ist.

Glatte Funktionen mit kompaktem Träger sind Schwartz-Funktionen, und die Gauss'schen Funktionen $e^{-a^2x^2}$, $a > 0$, sind ebenfalls Schwartz-Funktionen. Es ist klar, dass falls f eine Schwartz-Funktion ist, die Funktionen $x^\alpha D^\beta f$ wieder Schwartz-Funktionen sind. Ferner ist dann $x^\alpha D^\beta f$ stets integrierbar. Wir bezeichen den Vektorraum aller Schwartz-Funktionen mit $\mathcal{S}(\mathbb{R}^n)$. Weil $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ und weil $C_c^\infty(\mathbb{R}^n)$ dicht in $L^p(\mathbb{R}^n)$ liegt (für $p < \infty$), ist $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n 9$ ein dichter Unterraum.

Satz 4.11. *Sei f eine Schwartz-Funktion.*
- *Es gilt $\widehat{D_j f}(\xi) = i\xi_j \hat{f}(\xi)$. Insbesondere ist $\xi_j \hat{f}(\xi)$ beschränkt.*
- *$\hat{f}$ ist stetig differenzierbar und es gilt $\frac{\partial}{\partial \xi_j}\hat{f}(\xi) = -i\widehat{x_j f}(\xi)$.*
- *$\hat{f} \in \mathcal{S}(\mathbb{R}^n)$.*

Beweis. 1. Zunächst zeigen wir folgende Aussage: falls $g \in \mathcal{S}(\mathbb{R}^n)$, dann gilt
$$\int_{\mathbb{R}^n} D_j g(x)dx = 0.$$
Um dies einzusehen, dürfen wir ohne Beschränkung der Allgemeinheit annehmen, dass $j = 1$. Nach Fubini folgt
$$\int_{\mathbb{R}^n} D_j g(x)dx = \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} D_1 g(x,y)dxdy.$$
Das innere Integral ist aber $\lim_{R\to\infty} g(R,y) - g(-R,y) = 0$. Daher ist auch das Integral von $D_j g$ über $\mathbb{R}^n$ Null. Setze $g(x) = f(x)e^{-ix\xi}$. Es folgt
$$0 = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} D_j(f(x)e^{-ix\xi})dx =$$
$$= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} D_f(x)e^{-ix\xi}dx - \frac{1}{(2\pi)^{n/2}} i\xi_j \int_{\mathbb{R}^n} f(x)e^{-ix\xi}dx = \widehat{D_j f}(\xi) - i\xi_j \hat{f}(\xi).$$
2. Dies folgt durch Differentation unter dem Integral:
$$\frac{\partial}{\partial \xi_j}\hat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x)(-ix_j)e^{-i\xi x}dx = -i\widehat{x_j f}(\xi).$$
Die Voraussetzungen von Satz 2.34 sind erfüllt sind, weil f eine Schwartz-Funktion ist.

3. folgt durch Induktion aus den Teilen 1 und 2. □

Beispiel 4.12. *Betrachte die Funktion $f(x) = e^{-x^2/2}$. Wir wollen die Fouriertransformation dieser Funktion berechnen. Man beobachtet, dass $D_j f(x) + x_j f(x) = 0$. Es gilt also für die Fouriertransformation, nach Satz 4.11,*
$$0 = \widehat{D_j f}(\xi) + \widehat{x_j f}(\xi) = i\xi_j \hat{f}(\xi) + i\frac{\partial}{\partial \xi_j}\hat{f}(\xi).$$
Nun berechnen wir
$$\frac{\partial}{\partial \xi_j}(e^{\xi^2/2}\hat{f}(\xi)) = \xi_j e^{\xi^2/2}\hat{f}(\xi) - \xi_j e^{\xi^2/2}\hat{f}(\xi) = 0.$$
Also ist die Funktion $e^{\xi^2/2}\hat{f}(\xi)$ konstant, und es folgt $\hat{f}(\xi) = Ce^{-\xi^2/2}$ für eine Konstante $C = \hat{f}(0)$. Andererseits ist
$$\hat{f}(0) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-x^2/2}dx = 1$$

wegen 3.17.

Satz 4.13 (Fourier-Inversionsformel). *Sei $f \in \mathcal{S}(\mathbb{R}^n)$. Dann gilt:*

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi x} d\xi.$$

Beweis. Wir setzen die Definition der Fouriertransformation ein:

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi x} d\xi = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) e^{-iy\xi} e^{i\xi x} dy d\xi = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) e^{i\xi(x-y)} dy d\xi.$$

Leider ist der Integrand nicht in $L^1(\mathbb{R}^{2n})$ und wir brauchen einen Trick. Weil $\hat{f}$ eine L^1-Funktion ist, folgt aus dem Satz über dominierte Konvergenz

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi x} d\xi = \frac{1}{(2\pi)^{n/2}} \lim_{t \to 0} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-(t\xi)^2/2} e^{i\xi x} d\xi =$$

$$= \frac{1}{(2\pi)^n} \lim_{t \to 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) e^{-(t\xi)^2/2} e^{i\xi(x-y)} dy d\xi.$$

Nun ist der Integrand in L^1, und wir dürfen die Integrationsreihenfolge vertauschen, und erhalten

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi x} d\xi = \frac{1}{(2\pi)^n} \lim_{t \to 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) e^{-(t\xi)^2/2} e^{i\xi(x-y)} d\xi dy =$$

$$= \frac{1}{(2\pi)^n} \lim_{t \to 0} \int_{\mathbb{R}^n} (\int_{\mathbb{R}^n} e^{-(t\xi)^2/2} e^{i\xi(x-y)} d\xi) f(y) dy.$$

Das innere Integral ist

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-(t\xi)^2/2} e^{i\xi(x-y)} d\xi \overset{\xi=\eta/t}{=} \frac{1}{t^n} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-\eta^2/2} e^{-i\eta \frac{y-x}{t}} d\eta = \frac{1}{t^n} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-\eta^2/2} e^{-i\eta \frac{y-x}{t}} d\eta.$$

Mit der oben berechnenten Fouriertransformation von $e^{-x^2/2}$ ist dies gleich

$$\frac{1}{t^n} e^{-(\frac{y-x}{t})^2/2}.$$

In die obige Formel eingesetzt, erhalten wir

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi x} d\xi = \frac{1}{(2\pi)^{n/2}} \lim_{t \to 0} \int_{\mathbb{R}^n} \frac{1}{t^n} e^{-(\frac{y-x}{t})^2/2} f(y) dy$$

Wir substituieren $y = x + tz$ und erhalten

$$\frac{1}{(2\pi)^{n/2}} \lim_{t \to 0} \int_{\mathbb{R}^n} \frac{1}{t^n} e^{-(\frac{y-x}{t})^2/2} f(y) dy = \frac{1}{(2\pi)^{n/2}} \lim_{t \to 0} \int_{\mathbb{R}^n} e^{-z^2/2} f(x + tz) dz.$$

Weil f stetig und beschränkt ist, konvergiert der Integrand für $t \to 0$ dominiert gegen $e^{z^2/2} f(x)$, und dominierte Konvergenz sowie 3.17 zeigen

$$\frac{1}{(2\pi)^{n/2}} \lim_{t \to 0} \int_{\mathbb{R}^n} e^{-z^2/2} f(x + tz) dz = f(x)$$ □

Als Folgerung erhalten wir:

Korollar 4.14. *Die Fouriertransformation liefert einen linearen Isomorphismus $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ mit Inversem gegeben durch $g \mapsto \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} g(\xi) e^{i\xi x} d\xi$.*

Satz 4.15 (Plancherel-Formel). *Für eine Schwartz-Funktion f gilt*

$$\|\hat{f}\|_{L^2}^2 = \|f\|_{L^2}^2.$$

Beweis. Wir rechnen, mit einem ähnlichen Trick wie im Beweis von Satz 4.13

$$\|\hat{f}\|_2^2 = \int_{\mathbb{R}^n} \hat{f}(xi)\overline{\hat{f}(\xi)}d\xi = \lim_{t\to 0}\int_{\mathbb{R}^n} \hat{f}(xi)\overline{\hat{f}(\xi)}e^{-(t\xi)^2/2}d\xi =$$

$$= \frac{1}{(2\pi)^n}\lim_{t\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)e^{-i\xi x}\overline{f(y)}e^{i\xi y}e^{-(t\xi)^2/2}dxdyd\xi.$$

Weil f eine Schwartz-Funktion ist, ist der Integrand wegen des zusätzlichen Faktors eine L^1-Funktion auf $\mathbb{R}^{3n}$, und mit Fubini ist dieses Integral gleich

$$\frac{1}{(2\pi)^n}\lim_{t\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} e^{-i\xi(x-y)}e^{-(t\xi)^2/2}d\xi)f(x)\overline{f(y)}dxdy.$$

Im Beweis von Satz 4.13 haben wir das innere Integral ausgerechnet. Einsetzen liefert

$$\|\hat{f}\|_2^2 = \frac{1}{(2\pi)^n}\lim_{t\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{1}{t^n}e^{-(\frac{y-x}{t})^2/2}f(x)\overline{f(y)}dxdy =$$

(Substitution $y = x + tz$)

$$= \frac{1}{(2\pi)^n}\lim_{t\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-z^2/2}f(x)\overline{f(x+tz)}dxdz.$$

Weil f stetig ist, konvergiert der Integrand (dominiert) gegen $|f(x)|^2$, und aus dominierter Konvergenz folgt

$$\|\hat{f}\|_2^2 = \frac{1}{(2\pi)^n}\lim_{t\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-z^2/2}f(x)\overline{f(x+tz)}dxdz = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-z^2/2}f(x)\overline{f(x)}dxdz = \|f\|_{L^2}^2;$$

die letzte Gleichung folgt aus 3.17. □

5. Fourierreihen

5.1. Die Fourierreihe einer periodischen Funktion.

Definition 5.1. *Sei $T > 0$. Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt T-periodisch, falls $f(x+T) = f(x)$ für alle $x \in \mathbb{R}$ gilt.*

Beispiele solcher Funktionen sind natürlich die Funktionen

$$1,\ \cos(kx);\ \sin(kx);\ e^{ikx}\ (k \in \mathbb{Z}),$$

welche alle die Periode 2π haben. Die Bedeutung solcher Funktionen z.B. in der Physik liegt darin, dass sich periodische Vorgänge, z.B. Schwingungen sowie die Bewegung der Planeten durch periodische Vorgänge modellieren lassen. Wir wollen hier nur die Periode 2π betrachten; alle Ergebnisse lassen sich durch Reskalierung auf andere Periodenlängen verallgemeinern. Es stellt sich nun die grundlegende Frage, ob jede 2π-periodische Funktion als unendliche Reihe

$$f(x) = c_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + \sum_{k=1}^{\infty} b_k \sin(kx)$$

geschrieben werden kann und in welchem Sinne diese Reihe konvergiert (punktweise, punktweise fast überall, gleichmäßig, absolut, in L^p-Norm.....). Unser erstes Ziel ist es, heuristisch eine Formel für die Koeffizienten herzuleiten. Der Rechenaufwand wird *erheblich* reduziert, wenn man konsequent die komplexen Zahlen benutzt. Hierfür erinnern wir an die *Moivre-Formeln*

$$e^{ikx} = \cos(kx) + i\sin(kx);\ \cos(kx) = \frac{1}{2}(e^{ikx} + e^{-ikx});\ \sin(kx) = \frac{1}{2i}(e^{ikx} - e^{-ikx}).$$

Wir können damit $a_k \cos(kx) + b_k \sin(kx)$ auch als

$$a_k \cos(kx) + b_k \sin(kx) = \frac{1}{2}a_k(e^{ikx}+e^{-ikx}) + \frac{1}{2i}b_k(e^{ikx}-e^{-ikx}) = \frac{1}{2}(a_k - ib_k)e^{ikx} + \frac{1}{2}(a_k+ib_k)e^{-ikx}$$

ausdrücken. Setze

$$c_k := \frac{1}{2}(a_k - ib_k);\ c_{-k} := \frac{1}{2}(a_k + ib_k).$$

Damit können wie die obige unendliche Reihe in knapperer Form als

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \tag{5.2}$$

schreiben. Die Koeffizienten a_k und b_k werden durch die Formeln

$$b_k := ic_k - ic_{-k};\ a_k := c_k + c_{-k}$$

zurückgewonnen. *Außer wenn gewichtige Gründe dagegen sprechen*, werden wir immer die komplexe Form nehmen. Nun zur Frage, wie man aus der Reihendarstellung 5.2 die Koeffizienten c_k berechnen kann. Dafür zwei einfache Berechnungen.

Lemma 5.3. *Sei $f : \mathbb{R} \to \mathbb{C}$ T-periodisch und integrierbar. Dann gilt für jedes $x \in \mathbb{R}$*

$$\int_x^{x+T} f(t)dt = \int_0^T f(t)dt.$$

Beweis. Zunächst gilt wegen der Periodizität $\int_x^{x+T} f(t)dt = \int_{x+kT}^{x+(k+1)T} f(t)dt$ für jedes $k \in \mathbb{Z}$. Also können wir $0 \le x \le T$ annehmen und rechnen

$$\int_x^{x+T} f(t)dt = \int_x^T f(t)dt + \int_T^{T+x} f(t)dt = \int_x^T f(t)dt + \int_0^x f(t)dt = \int_0^T f(t)dt.$$

□

Als nächstes berechnen wir

$$\int_0^{2\pi} e^{ikt}e^{-ilt}dt = \int_0^{2\pi} e^{i(k-l)t}dt = \begin{cases} 2\pi & k = l, \\ \frac{1}{i(k-l)}[e^{i(k-l)t}]_0^{2\pi} = 0 & k \neq l, \end{cases} \tag{5.4}$$

und erhalten die bekannten *Orthogonalitätsrelationen*, welche später in der L^2-Theorie der Fourierreihen eine grundlegende Rolle spielen werden. Nun berechnen wir heuristisch, also ohne uns um Konvergenzfragen zu scheren,

$$\int_0^{2\pi} f(t)e^{-ikt}dt = \int_0^{2\pi} \sum_{n=-\infty}^{\infty} c_n e^{int}e^{-ikt}dt = \sum_{n=-\infty}^{\infty} c_n \int_0^{2\pi} e^{int}e^{-ikt}dt = 2\pi c_k.$$

Also sollte gelten

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}dt.$$

Dies motiviert die folgende Definition.

Definition 5.5. *Sei $f : \mathbb{R} \to \mathbb{C}$ 2π-periodisch und messbar, sowie $f \in L^1([0, 2\pi])$. Dann heißt*

$$\hat{f}(k) := \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(t)e^{-ikt}dt$$

der k-te Fourierkoeffizient *von f und die formale Reihe*

$$\sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\hat{f}(k)e^{ikx}$$

die Fourierreihe *von f. Der Grund für die seltsame Normierung wird später erklärt werden.*

Vereinbarung 2. *Wenn wir $\sum_{k=-\infty}^{\infty} a_k$ schreiben, so meinen wir* immer *die Folge*

$$\lim_{n\to\infty} \sum_{k=-n}^{n} a_k.$$

An diese Definition schließt sich folgender Fragenkatalog an:

(1) Unter welchen Bedingungen an f konvergiert die Fourierreihe? Hier kann "Konvergenz" viele Bedeutungen haben: punktweise, gleichmäßig, absolut, in L^p-Norm etc.
(2) Falls die Fourierreihe von f in irgendeinem Sinne konvergiert, welche Relation besteht zwischen der Grenzfunktion und der ursprünglichen Funktion f?

5.2. **Konvergenz im arithmetischen Mittel: Der Satz von Fejer und seine Konsequenzen.** Sei f eine periodische L^1-Funktion. Wir leiten zunächst eine Integralformel für die Partialsummen der Fourierreihe von f her. Setze

$$S_n f(x) := \sum_{k=-n}^{n} \frac{1}{\sqrt{2\pi}} \hat{f}(k) e^{ikx}.$$

Es gilt dann nach Definition

$$S_n f(x) = \frac{1}{2\pi} \int_0^{2\pi} \sum_{k=-n}^{n} f(t) e^{-ikt} e^{ikx} dt = \frac{1}{2\pi} \int_0^{2\pi} \sum_{k=-n}^{n} e^{ik(x-t)} f(t) dt = \int_0^{2\pi} D_n(x-t) f(t) dt,$$

wobei

$$D_n(y) := \frac{1}{2\pi} \sum_{k=-n}^{n} e^{iky}$$

der n*-te Dirichlet-Kern* ist. Eine explizite Formel leitet man mit der geometrischen Summenformel her. Zunächst sei $q = p^2 \in \mathbb{C} \setminus 0$.

$$\sum_{k=-n}^{n} q^k = q^{-n} \sum_{k=0}^{2n} q^k = q^{-n} \frac{1 - q^{2n+1}}{1-q} = \frac{q^{-n} - q^{n+1}}{1-q} =$$

$$\frac{p^{-2n} - p^{2n} p^2}{1-p^2} = p \frac{p^{-2n-1} - p^{2n+1}}{1-p^2} = \frac{p^{-2n-1} - p^{2n+1}}{p^{-1} - p} = \frac{p^{-1} q^{-n} - q^n p}{p^{-1} - p} =$$

$$= \frac{p^{-1} q^{-n}}{p^{-1} - p} - \frac{q^n p}{p^{-1} - p} = \frac{q^{-n}}{1-p^2} + \frac{q^n}{1-p^{-2}} = \frac{q^{-n}}{1-q} + \frac{q^n}{1-q^{-1}}.$$

Falls $q = e^{iy}$, $p = e^{\frac{1}{2} iy}$, so erhalten wir mit der Moivre-Formel $e^{ix} - e^{-ix} = 2i \sin(x)$

$$D_n(y) = \frac{1}{2\pi} \frac{\sin((n+\frac{1}{2})y)}{\sin(\frac{y}{2})}. \tag{5.6}$$

Wir sehen, mittels der Orthogonalitätsrelationen

$$\int_0^{2\pi} D_n(y) dy = 1.$$

Außerdem hat der Nenner die Nullstellen $2k\pi$, $k \in \mathbb{Z}$. An diesen Stellen ist D_n dennoch stetig, weil die Definition zeigt, dass D_n eine C^∞-Funktion ist. Es gilt

$$D_n(0) = \frac{2n+1}{2\pi}.$$

Wir zeigen nun eine schwächere Version der Konvergenz der Fourierreihen. Dafür zunächst eine einfache Beobachtung. Sei a_n eine Folge in einem normierten Vektorraum (z.B. $V = \mathbb{R}$). Wir nehmen an, dass $a_n \to a \in V$. Dann betrachten wir die Folge

$$A_n := \frac{1}{n+1} \sum_{k=0}^{n} a_k,$$

also die Folge der arithmetischen Mittel von a_k. Es gilt:

Lemma 5.7. *Falls $a_n \to a$, so gilt auch $A_n \to a$.*

Beweis. Sei $\epsilon > 0$. Wähle n_0, so dass $\|a - a_n\| \leq \epsilon/2$ für alle $n \geq n_0$. Dann gilt für $n \geq n_0$

$$\|A_n - a\| = \|\frac{1}{n+1}\sum_{k=0}^{n}(a-a_k)\| \leq \frac{1}{n+1}\sum_{k=0}^{n_0-1}\|a-a_k\| + \frac{1}{n+1}\sum_{n_0}^{n}\|a-a_k\| =: \frac{1}{n+1}C(\epsilon) + \frac{n-n_0}{n+1}\epsilon/2.$$

Für genügend große n ist dass $\leq \epsilon$. □

Es kann aber im allgemeinen durchaus gelten, dass die Folge der arithmetischen Mittel konvergiert, ohne dass die ursprüngliche Folge konvergiert. Ein Beispiel ist $a_n = (-1)^n$. In diesem Fall gilt

$$A_n := \frac{1}{n+1}\sum_{k=0}^{n}(-1)^k = \begin{cases} \frac{1}{n+1} & n\text{gerade} \\ 0 & n\text{ungerade.} \end{cases}$$

Also konvergiert A_n gegen 0. Der Prozess der arithmetischen Mittelung hat also durchaus die Chance, die Konvergenz einer Folge zu verbessern. In der Tat:

Satz 5.8 (Satz von Fejer). *Sei $f : \mathbb{R} \to \mathbb{C}$ stetig und 2π-periodisch. Dann konvergiert die Folge $F_n f(x) := \frac{1}{n+1}\sum_{k=0}^{n} S_n f(x)$ gleichmäßig gegen f.*

Für den Beweis schreiben wir zuerst

$$\frac{1}{n+1}\sum_{k=0}^{n} S_n f(x) = \frac{1}{n+1}\sum_{k=0}^{n}\int_0^{2\pi} D_n(x-t)f(t)dt = \int_0^{2\pi} F_n(x-t)f(t)dt,$$

wobei

$$F_n(y) := \frac{1}{n+1}\sum_{k=0}^{n} D_n(y)$$

der $n$te *Fejer-Kern* ist. Wir berechnen, wieder für $q = p^2 \in \mathbb{C} \setminus 0$:

$$\sum_{k=0}^{n}\sum_{m=-k}^{k} q^m = \sum_{k=0}^{n}\frac{q^{-k}}{1-q} + \frac{q^k}{1-q^{-1}} = \frac{1-q^{-n-1}+1-q^{n+1}}{(1-q)(1-q^{-1})} = \frac{(1-p^{-(n+1)})(1-p^{n+1})}{(1-p^2)(1-p^{-2})}.$$

Einsetzen $p = e^{iy/2}$ liefert

$$F_n(y) = \frac{1}{2\pi(n+1)}\frac{1-\cos((n+1)y)}{1-\cos(y)} \tag{5.9}$$

Diese Formel gilt nicht für $x = 2\pi k$. Entweder aus der Definition oder (einfacher) der L'Hopital-Regel erhalten wir aber

$$F_n(0) = \frac{n+1}{2\pi}.$$

Lemma 5.10. *Es gilt*

(1) *Die Funktion F_n ist 2π-periodisch und C^∞.*
(2) *Es gilt $\int_0^{2\pi} F_n(t)dt = 1$.*
(3) *Es gilt $F_n(y) \geq 0$ für alle n, für alle y.*
(4) *Für jedes $\eta > 0$ konvergiert die Funktionenfolge F_n auf dem Intervall $[\eta, 2\pi - \eta]$ gleichmäßig gegen 0.*

Beweis. (1) ist klar nach der Definition, und (2) folgt sofort aus der Formel $\int_0^{2\pi} D_n(t)dt = 1$. Für die anderen Eigenschaften nutzen wir die Formel (5.9). Hier ist (3) offensichtlich, denn $1-\cos(x) \in [0,2]$ für alle $x \in \mathbb{R}$. Ist $0 < \eta \leq x \leq 2\pi - \eta$, so sehen wir

$$0 \leq F_n(y) \leq \frac{1}{2\pi}\frac{1}{n+1}\frac{2}{1-\cos(x)} \leq \frac{1}{\pi}\frac{1}{n+1}\frac{1}{1-\cos(\eta)}. \tag{5.11}$$

Für $n \to \infty$ konvergiert das gleichmäßig gegen 0. □

Beweis von Satz 5.8. Der Beweis benutzt nur die Eigenschaften aus Lemma 5.10 und kann in allgemeinerem Kontext durchgezogen werden. Sei $f : \mathbb{R} \to \mathbb{C}$ stetig. Sei $\epsilon > 0$. Weil f periodisch und stetig ist, ist f gleichmäßig stetig. Es existiert also $\delta > 0$ so dass $|x-y| \leq 2\delta \Rightarrow |f(x)-f(y)| < \epsilon/2$. Wir rechnen

$$F_n f(x) - f(x) = \int_0^{2\pi} F_n(x-t)f(t)dt - f(x) \overset{5.10(2)}{=} \int_0^{2\pi} F_n(x-t)(f(t)-f(x))dt \overset{5.3}{=}$$

$$= \int_{x-\delta}^{x+\delta} F_n(x-t)(f(t)-f(x))dt + \int_{x+\delta}^{x+2\pi-\delta} F_n(x-t)(f(t)-f(x))dt =: I_1 + I_2.$$

Wir schätzen beide Integrale dem Betrag nach ab, und zwar wie folgt

$$|I_1| \leq \int_{x-\delta}^{x+\delta} |F_n(x-t)||f(t)-f(x)|dt \leq \frac{\epsilon}{2}\int_{x-\delta}^{x+\delta} |F_n(x-t)|dt \overset{5.10(3)}{=} \frac{\epsilon}{2}\int_{x-\delta}^{x+\delta} F_n(x-t)dt \overset{5.10(2)}{\leq} \frac{\epsilon}{2}$$

(dies gilt sogar für alle n!). Außerdem ist

$$|I_2| \leq \int_{x+\delta}^{x+2\pi-\delta} |F_n(x-t)||f(t)-f(x)|dt \leq 2\pi\frac{1}{\pi}\frac{1}{n+1}\frac{1}{1-\cos(\delta)}\|f\|_\infty = \frac{2}{n+1}\frac{1}{1-\cos(\delta)}\|f\|_\infty.$$

Also gilt: zu jedem $\epsilon > 0$ existiert ein $\delta > 0$, so dass für alle x und n gilt

$$|F_n f(x) - f(x)| \leq \frac{\epsilon}{2} + \frac{2}{n+1}\frac{1}{1-\cos(\delta)}\|f\|_\infty.$$

Daraus folgt aber die gleichmäßige Konvergenz. □

Der Satz von Fejer hat einige wichtige Konsequenzen.

Korollar 5.12. *Sei $f : \mathbb{R} \to \mathbb{C}$ stetig und 2π-periodisch. Dann gibt es zu jedem $\epsilon > 0$ ein trigonometrisches Polynom p, d.h. eine Funktion der Form*

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx},$$

so dass $\|f-p\|_\infty \leq \epsilon$. □

Satz 5.13 (Weierstrass'scher Approximationssatz)**.** *Sei $f : [a,b] \to \mathbb{C}$ stetig. Dann gibt es eine Folge p_n von Polynomen, welche auf $[a,b]$ gleichmäßig gegen f konvergiert. Falls f reellwertig ist, können reellwertige Polynome genommen werden.*

Beweis. Wir können nach Reskalierung annehmen, dass $[a,b] = [0,\pi]$. Des weiteren setzen wir f auf $[\pi, 2\pi]$ linear fort, so dass $f(2\pi) = f(0)$ und setzen f periodisch fort. Nach dem Satz von Fejer konvergiert die Folge der arithmetischen Mittel der Fourierreihe gleichmäßig gegen f. Wähle m mit $\|F_m f - f\|_\infty \leq \frac{1}{2n}$. Nun können wir schreiben

$$F_m f(x) = \sum_{k=-n}^{n} d_k e^{ikx}.$$

Aber die Potenzreihe $\sum_j (ik)^j \frac{1}{j!} x^j$ konvergiert gleichmäßig auf dem Intervall $[0, 2\pi]$ gegen e^{ikx} (Analysis I). Daraus baut man das entsprechende Polynom. □

Satz 5.14. *Sei $f \in L^p([0, 2\pi])$, $p \in [1, \infty)$. Dann konvergiert die Folge der arithmetischen Mittel $F_n f$ der Fourierreihe von f in der L^p-Norm gegen f.*

Beweis. Zunächst eine Abschätzung. Wir behaupten, dass $\|F_n f\|_p \leq \|f\|_p$ gilt. Dies folgt letztlich aus der Hölder-Ungleichung, aber mit einem Trick. Sei ganz abstrakt X ein Maßraum mit Maß ν, so dass $\nu(X) = 1$ gilt. Die Hölder-Ungleichung besagt dann

$$(\int_X |f| d\nu)^p \leq \|f\|_p^p \|1\|_q^p = \|f\|_p^p = \int_X |f|^p d\nu.$$

Sei nun X ein Maßraum mit Maß μ und g eine nichtnegative Funktion auf X mit $\int_X g d\mu = 1$. Dann ist $\nu(S) := \int_S g d\mu$ ein Maß mit $\nu(X) = 1$. Also sehen wir die etwas überraschende Ungleichung

$$(\int_X |f| g d\mu)^p \leq \int_X |f|^p g d\mu.$$

Dies können wir auf unser Problem anwenden, weil $F_n \geq 0$ und $\int_0^{2\pi} F_n = 1$ gilt. Es folgt

$$\|F_n f\|_p = \int_0^{2\pi} |\int_0^{2\pi} F_n(x-y) f(y) dy|^p dx \leq \int_0^{2\pi} \int_0^{2\pi} F_n(x-y) |f(y)|^p dy dx.$$

Vertausche die Integrationsreihenfolge, was nach Fubini erlaubt ist, denn $|f|^p$ ist eine L^1-Funktion. Wir erhalten

$$\|F_n f\|_p \leq \int_0^{2\pi} \int_0^{2\pi} F_n(x-y) |f(y)|^p dx dy = \|f\|_p^p.$$

Der Rest des Beweises ist einfach. Ist $f \in L^p$, so wähle, nach Satz 2.50, ein stetiges g mit kompaktem Träger in $(0, 2\pi)$ mit $\|g - f\|_p \leq \epsilon/3$ und setze g periodisch fort. Wir erhalten nun weil $\|F_n f\|_p \leq \|f\|_p$:

$$\|F_n f - f\|_p \leq \|F_n(f-g)\|_p + \|F_n g - g\|_p + \|g - f\|_p \leq \frac{2}{3}\epsilon + \|F_n g - g\|_p.$$

Weil $[0, 2\pi]$ endliches Maß hat, und weil $F_n g \to g$ gleichmäßig nach dem Satz von Fejer, konvergiert der letzte Summand gegen Null, wie behauptet. □

Satz 5.15 (Riemann-Lebesgue Lemma)**.** *Sei $f \in L^1([0, 2\pi])$. Dann ist die Folge $\hat{f}(k)$ der Fourierkoeffizienten von f eine Nullfolge.*

Beweis. Für jede L^1-Funktion f gilt

$$|\hat{f}(k)| = \frac{1}{\sqrt{2\pi}} |\int_0^{2\pi} f(t) e^{-ikt} dt| \leq \frac{1}{\sqrt{2\pi}} \|f\|_1.$$

Sei $\epsilon > 0$. Wähle ein trigonometrisches Polynom $g = \sum_{k=-n}^n a_k e^{ikx}$ mit $\|f - g\|_1 \leq \epsilon$. Es folgt $|\hat{f}(k) - \hat{g}(k)| \leq \epsilon$, aber auch $\hat{g}(k) = 0$ für $|k| > n$. □

Satz 5.16 (Eindeutigkeit der Fourierreihen)**.** *Es seien $f, g \in L^1([0, 2\pi])$ zwei periodische Funktionen mit gleichen Fourierkoeffizienten $\hat{f}(k) = \hat{g}(k)$. Dann gilt $f = g$ fast überall.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $g = 0$. Wir brauchen die Abschätzung

$$\|F_n f\|_1 \leq \|f\|_1. \tag{5.17}$$

Um dies zu zeigen, berechne

$$\|F_n f\|_1 = \int_0^{2\pi} |F_n f(x)| dx = \int_0^{2\pi} |\int_0^{2\pi} F_n(x-y) f(y) dy| dx \leq \int_0^{2\pi} \int_0^{2\pi} F_n(x-y)|f(y)| dy dx.$$

Nach Fubini dürfen wir die Reihenfolge der Integration vertauschen und erhalten

$$\int_0^{2\pi} \int_0^{2\pi} F_n(x-y)|f(y)| dy dx = \int_0^{2\pi} \int_0^{2\pi} F_n(x-y)|f(y)| dx dy = \int_0^{2\pi} |f(y)| dx dy = \|f\|_1.$$

Nun zum eigentlichen Beweis. Sei $\hat{f}(k) = 0$ für alle k. Dann gilt nach Definition $F_n f = 0$. Nun wähle eine stetige Funktion g mit $\|g - f\|_1 \leq \epsilon/3$. Dann gilt

$$\|f\|_1 = \|f - T_n f\|_1 \leq \|f - g\|_1 + \|g - T_n g\|_1 + \|T_n(g-f)\|_1 \leq \epsilon/3 + \|g - T_n g\|_1 + \epsilon/3,$$

wobei wir (5.17) verwendet haben. Nach dem Satz von Fejer gilt $\lim_n \|g - T_n g\|_1 = 0$, insgesamt also $\|f\|_1 \leq 1$. □

Satz 5.18 (Lokalisationssatz)**.** *Es sei f eine 2π-periodische L^1-Funktion. Sei $x \in [0, 2\pi]$ und $\delta > 0$ derart, dass $f|_{[x-\delta, x+\delta]}$ Lipschitz-stetig ist. Dann konvergiert die Fourierreihe von f auf $[x - \delta, x + \delta]$ gegen f.*

Dieser Satz hat die folgende, überraschende Konsequenz. Wenn f, g zwei L^1-Funktionen sind, die in einem Intervall $[x - \delta, x + \delta]$ übereinstimmen, so ist das Konvergenzverhalten der Fourierreihen dieser beiden Funktionen auf diesem Intervall gleich (obwohl die Fourierreihe von dem Verhalten der Funktion auf ganz $[0, 2\pi]$ abhängt.

Beweis. Es gilt

$$S_n f(x) - f(x) = \int_0^{2\pi} D_n(x-y)(f(y) - f(x)) dy = \int_0^{2\pi} D_n(t)(f(x+t) - f(x)) dt,$$

weil $\int_0^{2\pi} D_n(x-y) dy = 1$. Wir schreiben dieses Integral mit der Formel 5.6 als

$$\frac{1}{2\pi} \int_0^{2\pi} \cos(nt)(f(x+t) - f(x)) dt + \frac{1}{2\pi} \int_0^{2\pi} \sin(nt) \frac{\cos(t/2)(f(x+t) - f(x))}{\sin(t/2)} dt.$$

Der erste Summand ist der $n$te reelle Fourierkoeffizient der Funktion $t \mapsto f(x+t) - f(x)$ und konvergiert gegen Null wegen Satz 5.15. Wegen der Lipschitzbedingung gilt

$$|\frac{\cos(t/2)(f(x+t) - f(x))}{\sin(t/2)}| \leq \frac{L|t|}{|\sin(t/2)|}$$

für $|t| \leq \delta$ und für $|t| \geq \delta$ ist

$$|\frac{\cos(t/2)(f(x+t) - f(x))}{\sin(t/2)}| \leq \frac{1}{\sin(\delta/2}(|f(x)| + |f(t+x)|.$$

Insgesamt ist die Funktion $\frac{\cos(t/2)(f(x+t)-f(x))}{\sin(t/2)}$ integrierbar. Nach Satz 5.15 konvergiert also auch der zweite Summand gegen 0. □

5.3. **L^2-Theorie der Fourierreihen.** Die schönste Formulierung der Theorie der Fourierreihen ist im Kontext von L^2-Funktionen. Wir streben folgenden Satz an:

Satz 5.19 (L^2-Konvergenz der Fourierreihen)**.** *Sei $f \in L^2([0,2\pi])$. Dann konvergiert die Fourierreihe $S_n f$ von f in der L^2-Norm gegen f.*

Der Beweis macht entscheidenden Gebrauch von der Analogie des L^2-Skalarproduktes mit dem euklidischen Skalarprodukt auf $\mathbb{R}^n$.

Wir definieren $e_k \in L^2([0,2\pi])$ durch

$$e_k(x) := \frac{1}{\sqrt{2\pi}} e^{ikx}.$$

Die Orthogonalitätsrelationen (5.4) nehmen dann folgende Gestalt an

$$\langle e_k, e_l \rangle = \int_0^{2\pi} e_k(x)\overline{e_l(x)}dx = \frac{1}{2\pi}\int_0^{2\pi} e^{ikx}\overline{e^{ilx}}dx \frac{1}{2\pi}\int_0^{2\pi} e^{ikx}e^{-ilx}dx = \delta_{k,l}$$

(Kroneckersymbol). In Analogie mit der analytischen Geometrie sagen wir auch, dass $(e_k)_{k\in\mathbb{Z}}$ ein *Orthonormalsystem* in $L^2([0,2\pi])$ bildet.

Die Fourierkoeffizienten einer L^2-Funktion f haben die kompakte Form

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(t)e^{-ikt}dt = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(t)\overline{e^{ikt}}dt = \langle f, e_k\rangle$$

und die Fourierreihe ist

$$S_n f := \sum_{k=-n}^{n} \langle f, e_k\rangle e_k \in L^2([0,2\pi]). \tag{5.20}$$

Diese Formel sollte einem aus der Linearen Algebra bekannt vorkommen. Ist nämlich V ein endlichdimensionaler unitärer Vektorraum (z.B. $\mathbb{C}^n$) und $(e_1,\ldots,e_n)$ eine Orthonormalbasis, so gilt für jedes $v \in V$ die Formel

$$v = \sum_{i=1}^{n} \langle v, e_i\rangle e_i.$$

Die Fourierentwicklung ist also ein Analogon zur Entwicklung in eine Orthonormalbasis, aber im unendlich-dimensionalen Kontext.

Lemma 5.21 (Satz von Pythagoras)**.** *Seien $f_1,\ldots,f_n \in L^2([0,2\pi])$ paarweise orthogonal (d.h. $\langle f_i, f_j\rangle = 0$ für $i \neq j$). Dann gilt*

$$\|f_1 + \ldots + f_n\|^2 = \|f_1\|^2 + \ldots + \|f_n\|^2.$$

Beweis. Falls $n = 2$, so gilt

$$\|f_1+f_2\|^2 = \langle f_1+f_2, f_1+f_2\rangle = \langle f_1, f_1\rangle + \langle f_1, f_2\rangle + \langle f_2, f_1\rangle + \langle f_2, f_2\rangle = \|f_1\|^2 + \|f_2\|^2.$$

Für größere n argumentiert man durch Induktion über n. □

Als nächstes beobachten wir, dass

$$\langle f - S_n f, e_k\rangle = \begin{cases} 0 & n \geq |k| \\ \hat{f}_k & n < |k|. \end{cases}$$

Also steht $f - S_n f$ orthogonal auf allen e_k für $|k| \leq n$. Definiere $V_n := \mathrm{spann}\{e_{-n},\ldots,e_n\} \subset L^2([0,2\pi])$. Für $g \in V_n$ folgt aus der Bilinearität des Skalarproduktes

$$\langle f - S_n f, g\rangle = 0$$

und insbesondere weil $S_n f \in V_n$ auch

$$\langle f - S_n f, S_n f\rangle = 0.$$

Aus dem Satz von Pythagoras folgt ferner

$$\|f\|^2 = \|f - S_n f\|^2 + \|S_n f\|^2$$

und zusammen mit der Formel (5.20) ergibt sich (wieder mit Pythagoras)

$$\|f\|^2 = \|f - S_n f\|^2 + \sum_{k=-n}^{n} |\hat{f}(k)|^2. \tag{5.22}$$

Lemma 5.23 (Bestapproximationslemma)**.** *Für alle $g \in V_n$ gilt $\|f - g\| \geq \|f - S_n f\|$ und Gleichheit gilt genau dann, wenn $g = S_n f$. Mit anderen Worten: $S_n f$ ist die eindeutig bestimmte in der L^2-Norm am nächsten gelegene Funktion.*

Dieser Satz hat ein Analogon in der elementaren analytischen Geometrie (welches?).

Beweis. Schreibe $f - g = (f - S_n f) + (S_n f - g)$. Weil $S_n f - g \in V_n$ sind $(f - S_n f)$ und $(S_n f - g)$ orthogonal. Aus dem Satz von Pythagoras folgt also

$$\|f - g\|^2 = \|f - S_n f\|^2 + \|S_n f - g\|^2 \geq \|f - S_n f\|^2$$

und Gleichheit gilt genau dann, wenn $\|S_n f - g\| = 0$. □

Beweis von Satz 5.19. Sei $\epsilon > 0$. Nach dem 2. Approximationssatz 2.50 existiert eine stetige Funktion g mit kompaktem Träger in $(0, 2\pi)$ mit $\|f - g\|_2 \leq \epsilon/2$. Wir setzen g periodisch auf $\mathbb{R}$ fort. Aus dem Satz von Fejer 5.8 folgt, dass n_0 existiert mit $\|F_n g - g\|_\infty \leq \frac{1}{2\sqrt{2\pi}}\epsilon$ für $n \geq n_0$. Setze $q_n := F_n g$. Es gilt

$$\|q_n - g\|_2^2 = \int_0^{2\pi} |g(t) - q_n(t)|^2 dt \leq 2\pi \|g - q_n\|_\infty^2 \leq 2\pi \frac{1}{8\pi}\epsilon^2 = \frac{1}{4}\epsilon^2,$$

also $\|q - g_n\|_2 \leq \epsilon/2$ und zusammen $\|f - q_n\|_2 \leq \epsilon$.

Nun ist nach Konstruktion $q_n \in V_n$. Aus dem Bestapproximationslemma folgt dann

$$\|f - S_n f\| \leq \|f - q_n\| \leq \epsilon$$

für alle $n \geq n_0$, was zu zeigen war. □

Aus dem Satz 5.19 und der Gleichung 5.22 folgt sofort

Korollar 5.24 (Parzeval-Identität)**.** *Für jede Funktion $f \in L^2([0, 2\pi])$ gilt*

$$\|f\|_2^2 = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2.$$

Insbesondere gehört die Folge $(\hat{f}(k))_k$ der Fourierkoeffizienten zum Raum $L^2(\mathbb{Z})$.

Diese Sätze haben eine Umkehrung.

Proposition 5.25. *Sei $(a_k)_k \in L^2(\mathbb{Z})$ eine Folge. Dann gibt es genau ein Element $f \in L^2([0, 2\pi])$ mit $\hat{f}(k) = a_k$ für alle k.*

Beweis. Hier braucht man den Satz von Fischer-Riesz 2.47, welcher besagt, dass $L^2([0,2\pi])$ vollständig ist. Die Idee ist natürlich

$$f_n := \sum_{k=-n}^{n} a_k e_k$$

zu setzen und f als Grenzwert dieser Folge zu nehmen. In der Tat gilt für $n \geq m$

$$\|f_m - f_n\|^2 = \sum_{n<|k|\leq m} |a_k|^2 \|e_k\|^2 = \sum_{n<|k|\leq m} |a_k|^2$$

und daher ist $(f_n)_n$ eine L^2-Cauchyfolge und konvergiert gegen ein Element $f \in L^2([0,2\pi])$. Es gilt

$$\widehat{f_n}(k) = \begin{cases} a_k & n \geq |k| \\ 0 & n < |k| \end{cases}$$

nach Konstruktion und wegen

$$|\widehat{f}(k) - \widehat{f_n}| = |\langle f - f_n, e_k\rangle| \leq \|f - f_n\|_2 \|e_k\|_2 = \|f - f_n\|_2$$

gilt $\lim_{n\to\infty} \widehat{f_n}(k) = \widehat{f}(k)$, woraus die Behauptung folgt. □

Bemerkung 5.26. *Die Beweise waren im wesentlichen formaler Natur. Wir haben nur benutzt, dass $(e_k)_{k\in\mathbb{Z}}$ ein Orthonormalsystem ist und dass* $\operatorname{spann}\{e_k | k \in \mathbb{Z}\}$ *dicht in $L^2([0,2\pi])$ liegt (dies basierte auf der Regularität des Lebesgue-Maßes und dem Satz von Fejer). Aus diesem Grund lässt sich die obige Theorie sehr viel allgemeiner formulieren, nämlich als Orthonormalentwicklung in beliebigen Hilberträumen.*

Zum Abschluss des Kapitels über Fourierreihen noch ein Beispiel für die Macht dieser geometrischen Methoden.

Satz 5.27. *Es sei $f : \mathbb{R} \to \mathbb{C}$ eine stetig differenzierbare 2π-periodische Funktion. Dann konvergiert die Fourierreihe von f gleichmäßig und absolut gegen f.*

Beweis. Mit partieller Integration folgt

$$\hat{f'}(k) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f'(t) e^{-ikt} dt = ik \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t) e^{-ikt} dt = ik\hat{f}(k),$$

also $\hat{f}(k) = \frac{1}{ik}\hat{f'}(k)$. Nun ist die Folge der Fourierkoeffizienten von f' in $L^2(\mathbb{Z})$. Nach Cauchy-Schwarz folgt

$$\sum_{k=-\infty}^{k=\infty} |\hat{f}(k)| = |\hat{f}(0)| + \sum_{k\neq 0} |\frac{1}{k}\hat{f'}(k)| \leq |\hat{f}(0)| + (\sum_{k\neq 0} |\frac{1}{k}|^2)^{1/2} (\sum_{k\neq 0} |\hat{f'}(k)|^2)^{1/2} < \infty,$$

wobei benutzt wurde, dass die Reihe $\sum_{k\geq 1} \frac{1}{k^2}$ konvergiert. Daher konvergiert die Fourierreihe $\sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikx}$ gleichmäßig und absolut, gegen eine stetige Funktion g. Es folgt, dass $f = g$ fast überall, und somit $f = g$. □

6. Anfangsgründe der Funktionentheorie

6.1. **Holomorphe Funktionen.**

Definition 6.1. *Sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$, $z \in U$. Dann heißt f* komplex differenzierbar *in z, falls der Grenzwert*

$$f'(z) := \lim_{w \to z} \frac{f(w) - f(z)}{w - z} \in \mathbb{C}$$

existiert. f heißt holomorph *in U, falls f in jedem Punkt $z \in U$ komplex differenzierbar ist und falls die komplexe Ableitung $f' : U \to \mathbb{C}$ stetig ist.*

Bemerkung: man kann auf die Bedingung, dass f' stetig ist, verzichten, die Stetigkeit von f' *folgt*. Dies geschieht in den meisten Funktionentheoriebüchern. Wir wollen für die Entwicklung der Theorie jedoch die Methoden verwenden, die die allgemeine Integrationstheorie bietet, und es vereinfacht die Beweise erheblich, die Stetigkeit von f' vorauszusetzen.

Beispiel 6.2. *Polynome $f(z) = \sum_{k=0}^{n} a_k z^k$ sind holomorph, und es gilt $f'(z) = \sum_{k=1}^{n} k a_k z^{k-1}$. Summen, Produkte und (falls definiert) Quotienten und Kompositionen holomorpher Funktionen sind wieder holomorph, und die aus der Analysis I bekannten Ableitungsregeln übertragen sich auf die komplexe Ableitung.*

Beispiel 6.3. *Die Funktion $f(z) = \bar{z}$ ist* nicht *holomorph. Denn es gilt*

$$\frac{\bar{w} - \bar{z}}{w - z} = \begin{cases} 1 & w - z \in \mathbb{R} \\ -1 & w - z \in i\mathbb{R}. \end{cases}$$

Der Grenzwert kann nicht existieren. Ebensowenig ist $f(z) = |z|^2$ holomorph.

Zunächst wollen wir die Beziehung zwischen komplexer Differenzierbarkeit und der Ableitung Df im Sinne der Analysis II klären. Zunächst wiederholen wir die Definition der totalen Differenzierbarkeit. Seien V, W zwei endlich-dimensionale $\mathbb{R}$-Vektorräume, welche mit Normen ausgestattet seien. Sei $U \subset V$ offen und $f : U \to W$ sowie $x \in U$. Dann ist f differenzierbar in x, falls $Df(x) \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ existiert, so dass

$$\lim_{h \to 0} \frac{1}{\|h\|} \|f(x+h) - f(x) - Df(x)(h)\| = 0$$

gilt. In unserem Fall ist $V = W = \mathbb{C}$, und wir wählen den Absolutbetrag als Norm.

Satz 6.4. *Sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$. Dann sind äquivalent:*

1. *f ist komplex differenzierbar in z.*
2. *f ist differenzierbar in z (im Sinne der Analysis II), und die Ableitung $Df(z)$ ist durch Multiplikation mit einer komplexen Zahl $c \in \mathbb{C}$ gegeben.*

In diesem Falle ist $c = f'(z)$, das heißt, es gilt $Df(z)(h) = f'(z)h$ für alle $h \in \mathbb{C}$.

Beweis. $1 \Rightarrow 2$: Sei f komplex differenzierbar. Dann gilt

$$\lim_{h \to 0} \frac{1}{|h|} |f(z+h) - f(z) - f'(z)h| = \lim_{h \to 0} |\frac{f(z+h) - f(z)}{h} - f'(z)| = 0.$$

Also ist f differenzierbar, und die Ableitung ist durch $Df(z)(h) = f'(z)h$ gegeben.

$2 \Rightarrow 1$: Sei f differenzierbar und $c \in \mathbb{C}$ mit $Df(z)(h) = ch$. Es gilt dann

$$\lim_{h \to 0} |\frac{f(z+h) - f(z)}{h} - c| = \lim_{h \to 0} |\frac{f(z+h) - f(z) - ch}{h}| = \lim_{h \to 0} |\frac{1}{h}(f(z+h) - f(z) - Df(z)(h)| = 0,$$

also f komplex differenzierbar mit $f'(z) = c$. □

Die Aussage von Satz 6.4 kann, mit ein klein wenig Linearer Algebra, in eine weitere nützlichere Form gebracht werden.

Lemma 6.5. *Es sei $A : \mathbb{C} \to \mathbb{C}$ eine $\mathbb{R}$-lineare Abbildung. Dann sind äquivalent:*

(1) *Es gibt eine komplexe Zahl $c \in \mathbb{C}$ so dass $A(z) = cz$ für alle $z \in \mathbb{C}$.*
(2) *A ist $\mathbb{C}$-linear.*
(3) *Es gilt $A(i) = iA(1)$.*
(4) *Bezüglich der $\mathbb{R}$-Basis $\mathcal{B} = (1, i)$ von $\mathbb{C}$ hat A die Form*

$$M_{\mathcal{B}}(A) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathrm{Mat}_{2,2}(\mathbb{R}).$$

In diesem Fall gilt $c = a + ib = A(1)$.

Beweis. Die Implikationen $1 \Rightarrow 2 \Rightarrow 3$ sind trivial. $3 \Rightarrow 4$: man schreibe $A(1) = a + ib$. Dann gilt $A(i) = iA(1) = -b + ia$. Hieraus folgt die behauptete Form der darstellenden Matrix von A. $4 \Rightarrow 1$: Setze $c := a+ib$. Seien $x, y \in \mathbb{R}$ und $z = x+iy$. Man schreibe z in der Basis $\mathcal{B}$ aus. Nach den Regeln der Matrizenrechnung ist dann

$$A(z) = (ax - by) + i(bx + ay),$$

und weil

$$cz = (a + ib)(x + iy) = (ax - by) + i(bx + ay)$$

ist der Beweis vollständig. □

Wir wollen in Zukunft stets die implizite Notation $z = x + iy$ für komplexe Zahlen verwenden. Hier ist $x, y \in \mathbb{R}$, also $x = \Re(z)$ und $y = \Im(z)$. In diesem Sinne sind auch die partiellen Ableitungen $\frac{\partial}{\partial x} f$ und $\frac{\partial}{\partial y} f$ einer auf einer offenen Teilmenge $U \subset \mathbb{C}$ zu verstehen: wir fassen $f : U \to \mathbb{C}$ als Funktion zweier reeller Variabler auf, $f(x, y) = f(x+iy)$. Wir erinnern an den Zusammenhang der totalen Ableitung und den partiellen Ableitungen aus der Analysis II. Für Funktionen auf $U \subset \mathbb{C}$ gilt nämlich

$$\frac{\partial}{\partial x} f(z) = Df(z)(1); \ \frac{\partial}{\partial y} f(z) = Df(z)(i), \tag{6.6}$$

denn $\frac{\partial}{\partial x} f(z)$ ist die Richtungsableitung in Richtung des ersten Basisvektors, nämlich 1, und analog für $\frac{\partial}{\partial y} f(z)$.

Korollar 6.7. *Sei $U \subset \mathbb{C}$ offen, $f : U \to \mathbb{C}$, $z \in U$. Es sind äquivalent:*

(1) *f ist komplex differenzierbar bei z.*
(2) *f ist reell differenzierbar und $Df(z)$ ist $\mathbb{C}$-linear.*
(3) *f ist reell differenzierbar, und es gilt die* Cauchy-Riemann-Differentialgleichung

$$\frac{1}{2}(\frac{\partial}{\partial x} f(z) + i\frac{\partial}{\partial y} f(z)) = 0. \tag{6.8}$$

Unter diesen Bedingungen gilt $f'(z) = \frac{\partial}{\partial x} f(z) = \frac{1}{2}(\frac{\partial}{\partial x} f(z) - i\frac{\partial}{\partial y} f(z))$ und $Df(z)(h) = f'(z)h$.

Beweis. Die Äquivalenz 1 $\Leftrightarrow$ 2 folgt sofort aus Satz 6.4 und Lemma 6.5. 2 $\Leftrightarrow$ 3: Die Bedingung, dass $Df(z)$ eine $\mathbb{C}$-lineare Abbildung ist, ist nach Lemma 6.5 äquivalent zu der Bedingung $Df(z)(i) = iDf(z)(1)$, welche sich mit (6.6) auch als

$$\frac{\partial}{\partial y} f(z) = Df(z)(i) = iDf(z)(1) = i\frac{\partial}{\partial x} f(z)$$

ausdrücken lässt, also als

$$\frac{\partial}{\partial x} f(z) + i\frac{\partial}{\partial y} f(z) = 0,$$

und das ist gerade die Cauchy-Riemann-Differentialgleichung.

Die Formel für $f'(z)$ ist durch folgende Überlegung einzusehen. Falls $Df(z)$ komplex linear ist, ist es (Lemma 6.5) die Multiplikation mit der Zahl

$$f'(z) = Df(1) = \frac{\partial}{\partial x} f(z).$$

Schließlich ist, falls die Cauchy-Riemann-Differentialgleichung gilt

$$\frac{1}{2}(\frac{\partial}{\partial x} f(z) - i\frac{\partial}{\partial y} f(z)) = \frac{1}{2}(\frac{\partial}{\partial x} f(z) + \frac{\partial}{\partial x} f(z)) = f'(z).$$

Die Formel $Df(z)(h) = f'(z)h$ folgt sofort aus Lemma 6.5. □

In besonders kompakter Form lassen sich die Gleichungen mit dem sogenannten *Wirtinger-Kalkül* ausdrücken. Man definiert Differentialoperatoren (hier ist $f : U \to \mathbb{C}$ eine reell differenzierbare Funktion)

$$\frac{\partial}{\partial z} f(z) := \frac{1}{2}(\frac{\partial}{\partial x} f(z) - i\frac{\partial}{\partial y} f(z)); \; \frac{\partial}{\partial \bar{z}} f(z) := \frac{1}{2}(\frac{\partial}{\partial x} f(z) + i\frac{\partial}{\partial y} f(z)).$$

Die Cauchy-Riemann-Differentialgleichung lautet dann $\frac{\partial}{\partial \bar{z}} f(z) = 0$, und $f'(z) = \frac{\partial}{\partial z} f(z)$.

6.2. **Potenzreihen.** Wir betrachten nun *Potenzreihen* $\sum_{k=0}^{\infty} a_k(z-z_0)^k$, wobei $a_k \in \mathbb{C}$. Aus der Analysis I ist im wesentlichen der folgende Satz bekannt.

Satz 6.9. *Setze*

$$R := \frac{1}{\limsup_k |a_k|^{\frac{1}{k}}} \in [0, \infty]$$

(hier ist ausnahmsweise $\frac{1}{0} = \infty$ und $\frac{1}{\infty} = 0$ vereinbart). Dann gilt:

(1) *Falls $|z - z_0| > R$, so divergiert die Reihe $\sum_{k=0}^{\infty} a_k(z-z_0)^k$.*
(2) *Falls $|z - z_0| < R$, so konvergiert die Reihe $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ absolut.*
(3) *Auf jeder Kreisscheibe $B_r(z_0)$ mit Radius $r < R$ konvergiert die Reihe gleichmäßig.*

Die Zahl R heißt Konvergenzradius *der Potenzreihe.*

Beweis. Falls $|z - z_0| > R$ so gilt $\limsup_n |a_k|^{\frac{1}{k}}|z-z_0| > 1$, also ist

$$|a_k||z-z_0|^k \geq 1$$

für unendlich viele k. Somit ist die Reihe divergent. Falls $|z - z_0| < R$ so gilt $\limsup_n |a_k|^{\frac{1}{k}}|z-z_0| < q < 1$ für geeignetes q, also

$$|a_k||z-z_0|^k \leq q^k$$

für fast alle k, und aus dem Majorantenkriterium folgt absolute Konvergenz. Ferner folgt aus dieser Abschätzung auch die gleichmäßige Konvergenz auf jeder kompakten Kreisscheibe. $\square$

Wir können Potenzreihen gliedweise differenzieren, und die Grenzfunktion ist holomorph. Das folgt aus dem nächsten Satz.

Satz 6.10. *Es sei* $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ *eine Potenzreihe. Dann gilt:*

(1) *Die formal abgeleitete Potenzreihe* $\sum_{k=0}^{\infty} a_{k+1}(k+1)(z-z_0)^k$ *hat denselben Konvergenzradius wie die ursprüngliche Reihe.*
(2) *Sei* $R>0$ *der Konvergenzradius der Reihe und* $f(z) := \sum_{k=0}^{\infty} a_k(z-z_0)^k$ *die Grenzfunktion. Dann ist die Grenzfunktion* f *holomorph auf* $B_R(z_0)$*, und die Ableitung kann gliedweise berechnet werden:*
$$f'(z) = \sum_{k=0}^{\infty} a_{k+1}(k+1)(z-z_0)^k.$$
(3) *Insbesondere ist* f' *wieder holomorph, und* f *ist beliebig oft komplex differenzierbar.*

Beweis. 1: Die Reihe $\sum_{k=0}^{\infty}(k+1)a_{k+1}(z-z_0)^k$ ist genau dann konvergent, wenn $\sum_{k=0}^{\infty}(k+1)a_{k+1}(z-z_0)^{k+1} = \sum_{k=1}^{\infty} k a_k(z-z_0)^k$ es ist. Nun gilt
$$\limsup_k |a_k k|^{\frac{1}{k}} = \limsup_k |a_k|^{\frac{1}{k}} \lim k^{\frac{1}{k}} = \limsup_k |a_k|^{\frac{1}{k}}$$
denn $\lim k^{\frac{1}{k}} = 1$ (Analysis I).

2: wir wenden Satz 2.34 an, wobei wir $\mathbb{N}_0$ jetzt als Maßraum betrachten. Sei nämlich $g : \mathbb{N}_0 \times B_R(z_0) \to \mathbb{C}$ die Funktion $g(k,z) = a_k(z-z_0)^k$. Es gilt dann

(1) Für jedes $z \in B_R(z_0)$ ist $k \mapsto g(k,z)$ integrierbar über $\mathbb{N}_0$.
(2) Die Funktion $z \mapsto g(k,z)$ ist stetig differenzierbar.
(3) Zu jedem $w \in B_R(z_0)$ gibt es eine Umgebung $w \in U \subset B_R(z_0)$ und eine integrierbare Hilfsfunktion $h : \mathbb{N}_0 \to [0,\infty]$, so dass $|\frac{\partial}{\partial x} g(k,w)|, |\frac{\partial}{\partial y} g(k,w)| \leq h(k)$ für $w \in U$.

Die ersten beiden Aussagen sind klar. Für die dritte wähle $|w-z_0| < r < R$ und setze $U = B_r(z_0)$. Setze $h(k) := k|a_k|r^k$. Dann ist nach Satz 6.10 h integrierbar über $\mathbb{N}_0$, und es gilt
$$|\frac{\partial}{\partial y} g(k,w)| = |\frac{\partial}{\partial x} g(k,w)| = |g'(k,w)| = k|a_k||w|^k \leq h(k).$$
Nach Satz 2.34 ist die Grenzfunktion $f(z) := \sum_{k=0}^{\infty} a_k(z-z_0)^k = \int_{\mathbb{N}_0} g(k,z)d\mu(k)$ stetig differenzierbar, und wir dürfen Ableitung und "Integral" vertauschen. Insbesondere gilt
$$(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})f(z) = \sum_{k=0}^{\infty}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})(a_k(z-z_0)^k) = 0,$$
denn die Funktion $a_k(z-z_0)^k$ erfüllt die Cauchy-Riemannschen Differenzialgleichungen 6.8. Folglich erfüllt auch f 6.8 und ist daher holomorph. Die Formel für die Ableitung von f ergibt sich ebenso. $\square$

6.3. **Kurvenintegrale und der Cauchysche Integralsatz.** In Anlehnung an die Analysis I formulieren wir nun den Begriff der *Stammfunktion*.

Definition 6.11. *Es sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$ stetig. Eine* Stammfunktion *von f ist eine holomorphe Funktion $F : U \to \mathbb{C}$ mit $F' = f$.*

Anders als in der Analysis I ist die Existenz einer Stammfunktion keineswegs garantiert und die Antwort führt auf komplett unerwartete Phänomene. Erstens stellt sich heraus, dass die Holomorphie von f eine notwendige Bedingung ist, und zweitens hängt die Antwort von der Geometrie von U ab. Doch der Reihe nach. Wir beginnen mit einem neuen, anschaulich einleuchtenden Begriff.

Definition 6.12. *Sei $U \subset \mathbb{C}$ offen und $z_0, z_1 \in U$. Ein* Weg von z_0 nach z_1 *ist eine stetige Abbildung $c : [0,1] \to U$ mit $c(i) = z_i$. $U \subset \mathbb{C}$ heißt* wegzusammenhängend*, falls zu $z_0, z_1 \in U$ ein Weg von z_0 nach z_1 existiert. (Allgemeiner kann man diese Begriffe für metrische Räume X anstatt von $U \subset \mathbb{C}$ formulieren).*

Um zu sehen, wie eine Stammfunktion F aus der gegebenen Funktion f konstruiert werden kann, nehmen wir zunächst an, dass eine Stammfunktion existiert. Sei also $F : U \to \mathbb{C}$ eine Stammfunktion (d.h. $F' = f$ und sei $z_0 \in U$ fest und sei $z \in U$ ein weiterer Punkt sowie $c : [a,b] \to U$ ein C^1-Weg von z_0 nach z_1. Es gilt dann

$$F(z)-F(z_0) = F(c(1))-F(c(0)) = \int_a^b \frac{d}{dt}F(c(t))dt = \int_a^b F'(c(t))\dot{c}(t)dt = \int_a^b f(c(t))\dot{c}(t)dt,$$

wie aus dem Hauptsatz und der Kettenregel folgt. Aus der Kettenregel aus Analysis II folgt nämlich

$$\frac{d}{dt}F(c(t)) = Df(c(t))(Dc(t)) = f'(c(t))c'(t),$$

(in der zweiten Gleichung wurde die Holomorphie von f wesentlich benutzt, siehe Korollar 6.7!).

Korollar 6.13. *Sei $U \subset \mathbb{C}$ wegzusammenhängend und offen und $f : U \to \mathbb{C}$. Dann unterscheiden sich zwei Stammfunktionen von f durch Addition einer Konstante.*

Beweis. Seien F, G zwei Stammfunktionen von f. Dann ist $F - G$ eine Stammfunktion von 0 und wir müssen zeigen, dass $H = F - G$ konstant ist. Sei $z \in U$ und wähle einen C^1-Weg $c : [0,1] \to U$ von z_0 nach z (die Definition des Wegzusammenhanges liefert zunächst nur die Existenz eines stetigen Weges; das man C^1-Wege wählen kann, folgt aus Satz 6.25). Es gilt dann

$$H(z) - H(z_0) = \int_0^1 H'(c(t))\dot{c}(t)dt = 0. \qquad \square$$

Diese Beobachtung ermuntert zu folgender Definition.

Definition 6.14. *Es sei $f : U \to \mathbb{C}$ stetig und $c : [a,b] \to U$ ein C^1-Weg ($a \leq b$). Das* komplexe Kurvenintegral *von f über c ist definiert als*

$$\int_c f(z)dz = \int_a^b f(c(t))\dot{c}(t)dt.$$

Beispiel 6.15. *Sei $U = \mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ und $f(z) = z^k$, $k \in \mathbb{Z}$, sowie $c : [0,1] \to U$, $c(t) = e^{2\pi i t}$. Dann gilt*

$$\int_c f(z)dz = \int_0^1 (e^{2\pi it})^k 2\pi i e^{2\pi it} dt = 2\pi i \int_0^1 e^{2\pi i(k+1)t} dt = \begin{cases} 2\pi i & k = -1 \\ 0 & k \neq -1. \end{cases}$$

Die Funktionen $f(z) = z^k$ für $k \neq -1$ haben Stammfunktionen, zum Beispiel $F(z) = \frac{1}{k+1} z^{k+1}$. Hingegen hat $f(z) = \frac{1}{z}$ keine *auf ganz $\mathbb{C}^\times$ definierte Stammfunktion, denn sonst müsste $\int_c f(z)dz = F(c(1)) - F(c(0)) = 0$ gelten. Dieses Phänomen ist der Kern der gesamten Funktionentheorie und wird später eingehender untersucht werden.*

Wir würden nun gerne eine Stammfunktion F durch die Festlegung $F(z_0) = 0$ und

$$F(z) := \int_{c_z} f(z)dz$$

definieren, wobei c_z ein Weg von z_0 nach z ist. Das offensichliche Problem ist, ob dies von der Wahl des Weges abhängt, und diese Bedingung ist notwendig für die Existenz einer Stammfunktion. Wir nähern uns dieser Frage, indem wir den Weg c *deformieren*, mit anderen Worten, wir betrachten eine C^2-Abbildung $h : [0,1] \times [a,b] \to U$. Für $s \in [0,1]$ ist dann $h_s(t) := h(s,t)$ ein Weg von $h_s(a)$ nach $h_s(b)$.

Satz 6.16 (Homotopieformel)**.** *Sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$ holomorph, sowie $h : [0,1] \times [a,b] \to U$ eine C^2-Abbildung. Dann ist $G : [0,1] \to \mathbb{C}$, $G(s) := \int_{h_s} f(z)dz$ stetig differenzierbar, und es gilt*

$$G'(s) = [f(h(s,t)) \frac{\partial}{\partial s} h(s,t)]_{t=a}^{t=b}.$$

Beweis. Die Definition des Kurvenintegrals zeigt

$$G(s) = \int_a^b f(h(s,t)) \frac{\partial}{\partial t} h(s,t) dt.$$

Aus Satz 2.34 folgt, dass G stetig differenzierbar ist und dass wir unter dem Integral differenzieren dürfen. Es gilt

$$G'(s) = \int_a^b \frac{\partial}{\partial s}(f(h(s,t)) \frac{\partial}{\partial t} h(s,t)) dt = \int_a^b \frac{\partial}{\partial s} f(h(s,t)) \frac{\partial}{\partial t} h(s,t) dt + \int_a^b f(h(s,t)) \frac{\partial}{\partial s} \frac{\partial}{\partial t} h(s,t) dt. \tag{6.17}$$

Der zweite Summand berechnet sich mit dem Satz von Schwarz sowie mit partieller Integration:

$$\int_a^b f(h(s,t)) \frac{\partial}{\partial s} \frac{\partial}{\partial t} h(s,t) dt = \int_a^b f(h(s,t)) \frac{\partial}{\partial t} \frac{\partial}{\partial s} h(s,t) dt =$$

$$= [f(h(s,t)) \frac{\partial}{\partial s} h(s,t)]_{t=a}^{t=b} - \int_a^b \frac{\partial}{\partial t} f(h(s,t)) \frac{\partial}{\partial s} h(s,t) dt. \tag{6.18}$$

Wegen der Holomorphie von f (!!) gilt

$$\frac{\partial}{\partial t} f(h(s,t)) \frac{\partial}{\partial s} h(s,t) = f'(h(s,t)) \frac{\partial}{\partial t} h(s,t) \frac{\partial}{\partial s} h(s,t) = \frac{\partial}{\partial s} f(h(s,t)) \frac{\partial}{\partial t} h(s,t). \tag{6.19}$$

Mit diesem Symmetrietrick ergibt sich durch Einsetzen sofort:

$$G'(s) = [f(h(s,t))\frac{\partial}{\partial s}h(s,t)]_{t=a}^{t=b}. \qquad \square$$

Um die Bedeutung des letzten Satzes zu erkennen, formulieren wir einen neuen Begriff (dies ist einer der Grundbegriffe der *algebraischen Topologie*).

Definition 6.20. *Sei $U \subset \mathbb{C}$ offen. Eine* Homotopie von Wegen in U *ist eine stetige Abbildung $h : [0,1] \times [a,b] \to U$. Wir fassen h als Familie von Wegen $(h_s)_{s\in[0,1]}$, $h_s(t) = h(s,t)$ auf. Eine* Homotopie von Wegen mit festen Endpunkten *ist eine Homotopie h mit der zusätzlichen Eigenschaft, dass $h_s(a) = h_0(a)$ und $h_s(b) = h_0(b)$ für alle $s \in [0,1]$ gilt. Eine* Homotopie von geschlossenen Wegen *ist eine Homotopie von Wegen mit der zusätzlichen Eigenschaft, dass $h_s(a) = h_s(b)$ für alle $s \in [0,1]$ gilt. Eine* Nullhomotopie *eines geschlossenen Weges c ist eine Homotopie geschlossener Wege h mit $h_0 = c$ und $h_1 = const$.*

Wieder lässt sich das alles allgemein in metrischen Räumen formulieren. Man veranschauliche sich diese Vokabeln durch Bilder.

Korollar 6.21 (Homotopieinvarianz des Integrals holomorpher Funktionen)**.** *Es sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$ holomorph. Ferner sei $h : [0,1] \times [a,b] \to U$ eine C^2-Abbildung. Falls h eine Homotopie mit festen Endpunkten ist oder eine Homotopie von geschlossenen Wegen, so ist $\int_{h_s} f(z)dz$ unabhängig von s.*

Falls h eine Nullhomotopie ist, so gilt $\int_{h_0} f(z)dz = 0$.

Beweis. Sei G wie in Satz 6.16 definiert. Es gilt dann

$$G'(s) = [f(h(s,t))\frac{\partial}{\partial s}h(s,t)]_{t=a}^{t=b} = f(h(s,b))\frac{\partial}{\partial s}h(s,b) - f(h(s,a))\frac{\partial}{\partial s}h(s,a).$$

Wenn h eine Homotopie mit festen Endpunkten ist, so gilt $\frac{\partial}{\partial s}h(s,a) = \frac{\partial}{\partial s}h(s,b) = 0$. Wenn h eine Homotopie geschlossener Wege ist, so ist $\frac{\partial}{\partial s}h(s,a) = \frac{\partial}{\partial s}h(s,b)$ und $f(h(s,b)) = f(h(s,a))$ und in beiden Fällen heben sich die beiden Summanden weg. Für die letzte Aussage beobachtet man noch, dass das Kurvenintegral über einen konstanten Weg stets Null ist. $\square$

Korollar 6.22 (Cauchy'scher Integralsatz)**.** *Sei U sternförmig und $f : U \to \mathbb{C}$ holomorph. Dann gilt für jeden geschlossenen C^2-Weg c in U: $\int_c f(z)dz = 0$.*

Zur Erinnerung: eine Menge $U \subset \mathbb{C}$ heißt sternförmig bezüglich z_0, falls für alle $z \in U$ und alle $t \in [0,1]$ gilt: $(1-t)z_0 + tz \in U$. Es ist klar, dass sternförmige Mengen wegzusammenhängend sind.

Beweis. Sei $z_0 \in U$ ein Punkt, um den U sternförmig ist und $c : [a,b] \to U$ ein geschlossener Weg. Für $s \in [0,1]$ setze $c_s(t) = sz_0 + (1-s)c(t)$. Dies ist eine Homotopie geschlossener Wege. Es folgt

$$\int_c f(z)dz = \int_{c_0} f(z)dz = \int_{c_1} f(z)dz.$$

Der Weg c_1 ist aber konstant, und daher gilt $\int_{c_1} f(z)dz = \int_a^b f(c_1(t))c_1'(t)dt = 0$. $\square$

Korollar 6.23. *Sei U offen und sternförmig. Sei $f : U \to \mathbb{C}$ holomorph. Dann besitzt f eine Stammfunktion.*

Beweis. Sei $c_z(t) := (1-t)z_0 + tz$. Dies ist ein Weg in U von z_0 nach z. Setze $F(z) := \int_{c_z} f(z)dz$. Es gilt dann

$$F(z) = \int_0^1 f((1-t)z_0 + tz)(z - z_0)dt$$

nach der Definition des Kurvenintegrals. Sei $w \in \mathbb{C}$ nahe bei 0. Definiere eine Homotopie

$$h(s,t) := (1-t)z_0 + tz + tsw.$$

Dann gilt $h(s,t) = c_{z+sw}(t)$, und das liegt für alle $t \in [0,1]$, wenn nur $|w|$ klein genug ist. Aus Satz 6.16 folgt

$$\frac{d}{ds}F(z+sw) = [f(h(s,t))\frac{\partial}{\partial s}h(s,t)]_{t=0}^{t=1} = f(h(s,1))\frac{\partial}{\partial s}h(s,1) = f(z+sw)w.$$

Hieraus sehen wir, dass

$$\frac{\partial}{\partial x}F(z) = \frac{d}{ds}|_{s=0}F(z+s) = f(z)$$

und

$$\frac{\partial}{\partial y}F(z) = \frac{d}{ds}|_{s=0}F(z+si) = f(z)i$$

gilt. Also erfüllt F die Cauchy-Riemann-Differentialgleichung, und F ist holomorph. Ferner ist $F'(z) = f(z)$. □

6.4. **Exponentialfunktion und Logarithmus im Komplexen.** Wie im Reellen definieren wir die komplexe Exponentialfunktion durch die Potenzreihe

$$\exp(z) := e^z := \sum_{k=0}^{\infty} \frac{1}{k!}z^k.$$

Den Konvergenzradius dieser Reihe bestimmt man am besten mit dem Quotientenkriterium, welches zeigt, dass die Reihe für alle $z \in \mathbb{C}$ konvergiert, also ist der Konvergenzradius ∞. Durch gliedweise Differentation erhalten wir

$$\frac{d}{dz}e^z = e^z.$$

Wie in Analysis I zeigt man die Gleichung $e^{z+w} = e^z e^w$: Zunächst gilt

$$\frac{d}{dz}(e^{-z}e^z) = 0,$$

woraus $e^{-z}e^z = 1$ folgt. Ferner gilt

$$\frac{d}{dz}(e^{z+w}e^{-z}e^{-w}) = 0$$

und somit $e^{z+w} = e^z e^w$.

Wir definieren außerdem

$$\cos(z) = \frac{1}{2}(e^{iz} + e^{-iz});\ \sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz}).$$

Für reelle Argumente stimmen diese Funktionen mit den schon bekannten Funktionen überein. Ferner gilt die Moivre-Formel

$$e^{iz} = \cos(z) + i\sin(z)$$

auch für komplexe Argumente. Falls $x, y \in \mathbb{R}$, so sehen wir

$$e^{x+iy} = e^x(\cos(y) + i\sin(y))$$

und daraus

$$|e^{x+iy}| = e^x \sqrt{\cos(y)^2 + \sin(y)^2} = e^x.$$

Korollar 6.24. *Die komplexe Exponentialfunktion ist ein Homomorphismus von Gruppen* $(\mathbb{C}, +) \to (\mathbb{C}^\times, \cdot)$*, und es gilt* $\ker(\exp) = 2\pi i\mathbb{Z}$.

Beweis. Falls $e^z = 1$, so gilt $1 = |e^z| = e^{\Re(z)}$, also $\Re(z) = 0$. Aber $e^{iy} = 1$ impliziert $y \in 2\pi\mathbb{Z}$. $\square$

Sei nun $\mathbb{C}^- = \mathbb{C} \setminus \{z | z \geq 0\}$. Dies ist sternförmig um 1. Die Funktion $f(z) = \frac{1}{z}$ ist holomorph und es gibt genau eine Stammfunktion $\mathrm{Log} : \mathbb{C}^- \to \mathbb{C}$ mit $\mathrm{Log}(1) = 0$.

Für $z > 0$ gilt $F(z) = \log(z)$ (dies sieht man am einfachsten, indem man das Integral in Korollar 6.23 direkt ausrechnet). Also ist Log eine Erweiterung des reellen Logarithmus. Berechne

$$\frac{d}{dz}(\frac{1}{z} e^{\mathrm{Log}(z)}) = -\frac{1}{z^2} e^{\mathrm{Log}(z)} + \frac{1}{z} \mathrm{Log}'(z) e^{\mathrm{Log}(z)} = 0.$$

Hieraus folgt, dass $\frac{1}{z} e^{\mathrm{Log}(z)}$ konstant ist, und durch Einsetzen von $z = 1$ findet man $\frac{1}{z} e^{\mathrm{Log}(z)} = 1$, also

$$\exp(\mathrm{Log}(z)) = z.$$

Mit der anderen Gleichung $\mathrm{Log}(e^y) = y$*, die man von einem Logarithmus erwartet muss man aufpassen. Diese gilt nicht immer.* Diese Funktion Log heißt *Hauptzweig des Logarithmus*. Wir treffen an dieser Stelle auf ein sehr interessantes Phänomen. Sei $k \in \mathbb{Z}$. Die Funktion $2\pi i k + \mathrm{Log}(z)$ ist eine weitere Stammfunktion von $\frac{1}{z}$, aber es gilt auch

$$\exp(2\pi i k + \mathrm{Log}(z)) = \exp(2\pi i k) \exp(\mathrm{Log}(z)) = z.$$

Man könnte also mit Fug und Recht auch $2\pi i k + \mathrm{Log}(z)$ als Logarithmus ansehen, und nicht viel spricht dafür, der Funktion $\mathrm{Log}(z)$ den Vorzug zu geben (außer, dass der Logarithmus aus Analysis fortgesetzt wird). Wie sind die verschiedenen "Zweige" des Logarithmus miteinander verbunden? Nun, betrachte den Weg $c : (-\pi, \pi) \to \mathbb{C}^-$, $c(t) = e^{it}$. Wir berechen

$$\frac{d}{dt} \mathrm{Log}(c(t)) = \frac{c'(t)}{c(t)} = i,$$

und daraus folgt wegen $\mathrm{Log}(c(0)) = 0$ $\mathrm{Log}(c(t)) = it$ und

$$\lim_{t \to \pm\pi} \mathrm{Log}(c(t)) = \pm\pi.$$

Daraus folgt, dass $\mathrm{Log}(z)$ nicht stetig und schon gar nicht holomorph auf $\mathbb{C}^\times$ fortgesetzt werden kann.

Mit Hilfe des komplexen Logarithmus kann man allgemeine Potenzen definieren:

$$z^s := e^{s \,\mathrm{Log}(z)}.$$

Hier ist wieder $z \in \mathbb{C}^-$ zu setzen, und falls $s \notin \mathbb{Z}$, so findet man ähnliche Unstetigkeitsphänomene an der negativen reellen Achse.

6.5. **Verallgemeinerung auf stückweise C^2-Wege.** Dieses Kapitel kann beim ersten Lesen ohne Verlust überschlagen werden. Zunächst liefern wir den Beweis nach, dass in wegzusammenhängenden offenen Mengen $U \subset \mathbb{C}$ je zwei Punkte durch einen glatten Weg verbunden werden können.

Satz 6.25. *Sei $U \subset \mathbb{C}$ offen. Dann ist U genau dann wegzusammenhängend, wenn zu je zwei $z_0, z_1 \in U$ ein beliebig oft stetig differenzierbarer Weg $c : [0,1] \to U$ mit $c(i) = z_i$ existiert (hier könnte man $\mathbb{R}^n$ statt $\mathbb{C}$ haben).*

Beweis. Es ist zu zeigen, dass falls U wegzusammenhängend ist und falls $z_0, z_1 \in U$, eine C^∞-Abbildung $c : [0,1] \to U$ existiert mit $c(i) = z_i$. Sei $b : [0,1] \to U$ ein Weg von z_0 nach z_1. Sei $\epsilon := \inf_t \operatorname{dist}(b(t), U^c) > 0$ (Satz 1.29). Nach dem Weierstraßschen Approximationssatz gibt es ein Polynom p, so dass $|p(t) - b(t)| \leq \epsilon/4$ für alle t gilt. Dann ist

$$p(t) - ((1-t)(p(0) - c(0)) + t(p(1) - c(1))$$

ein glatter Weg von z_0 nach z_1, und es gilt

$$|p(t) - ((1-t)(p(0) - c(0)) + t(p(1) - c(1)) - c(t)| \leq \frac{3}{4}\epsilon,$$

also ein Weg in U. ☐

Definition 6.26. *Ein stetiger Weg $c : [a,b] \to U$ heißt* stückweise C^2*, falls eine Unterteilung $a = x_0 < x_1 < \ldots < x_r = b$ existiert, so dass $c|_{[x_{j-1},x_j]}$ ein C^2-Weg ist. Für eine stetige Funktion $f : U \to \mathbb{C}$ definieren wir das Kurvenintegral durch die Formel*

$$\int_c f(z)dz = \sum_{j=1}^{r} \int_{c|_{[x_{j-1},x_j]}} f(z)dz.$$

Wir wollen nun Korollar 6.21 verallgemeinern.

Proposition 6.27. *Es sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$ holomorph. Ferner sei $h : [0,1] \times [a,b] \to U$ eine stetige Homotopie, und h_0, h_1 seien stückweise C^2-Wege. Falls h eine Homotopie mit festen Endpunkten ist oder eine Homotopie von geschlossenen Wegen, so ist $\int_{h_s} f(z)dz$ unabhängig von s.*

Der Cauchy-Integralsatz lässt sich ebenfalls verallgemeinern, dahingehend, dass der geschlossene Weg c bloß stückweise C^2 ist.

Lemma 6.28. *Sei $U \subset \mathbb{C}$ offen und $h : [0,1] \times [a,b] \to U$ eine stetige Homotopie. Seien h_0 und h_1 stückweise C^2. Dann gibt es eine Folge von glatten Wegen $h_0 = g_0, g_1, \ldots, g_r = h_1$ in U, so dass die Homotopie $(s,t) \mapsto (1-s)g_{j-1}(t) + sg_j(t)$ ganz in U liegt. Falls h eine Homotopie mit festen Endpunkten ist, so können wir $g_j(a) = h_0(a)$, $g_j(b) = h_0(b)$ wählen, und falls h eine Homotopie geschlossener Wege ist, können wir $g_j(a) = g_j(b)$ wählen.*

Beweis. Wir führen den Beweis für feste Endpunkte durch, der andere Fall ist analog. Seien z_0, z_1 die Endpunkte. Sei $\epsilon := \inf_{s,t} \operatorname{dist}(h(s,t), U^c) > 0$. Weil h gleichmäßig stetig ist, gibt es $r \in \mathbb{N}$, so dass $|s - s'| \leq 1/r \Rightarrow |h(s,t) - h(s',t)| \leq \epsilon/5$. Nach dem Weierstraßschen Approximationssatz (und dem Argument aus dem Beweis von Satz 6.25) gibt es für $j = 1, \ldots, r-1$ glatte Wege $g_j : [0,1] \to U$ von z_0 nach z_1 mit $|g_j(t) - h(\frac{j}{r}, t)| \leq \epsilon/5$ für alle t. Setze $g_0 := c_0$ und $g_r := c_1$. Es gilt dann $|g_j(t) - g_{j-1}(t)| \leq \frac{3}{5}\epsilon$ und $\operatorname{dist}(g_j(t), U^c) \geq \frac{4}{5}\epsilon$. Daraus folgt, dass die lineare Verbindungsstrecke von g_{j-1} nach g_j ganz in U liegt. ☐

Beweis von Proposition 6.27. Weil die in Lemma 6.28 konstruierten Wege g_{j-1}, g_j homotop sind (durch eine Homotopie mit festen Endpunkten oder eine Homotopie geschlossener Wege, je nachdem), für $j = 2, \dots r-1$, folgt

$$\int_{g_{j-1}} f(z)dz = \int_{g_j} f(z)dz$$

für $j = 2, \dots, r-1$. Dasselbe Argument wendet man an, um $\int_{h_0} f(z)dz = \int_{g_1} f(z)dz$ und $\int_{g_{r-1}} f(z)dz = \int_{h_1} f(z)dz$ zu zeigen, nur dass man das Integrationsintervall in glatte Stücke zerlegen muss. Man wendet dann den Satz 6.16 auf jedes Teilintervall an. Die Details sind mühsam zu notieren, aber ohne weitere Ideen und unterbleiben hier. □

6.6. **Die Cauchy'sche Integralformel und ihre Konsequenzen.** Wir haben nun das Tor aufgestoßen zu einer Reihe phantastischer Sätze.

Satz 6.29 (Cauchy'sche Integralformel)**.** *Es sei $f : U \to \mathbb{C}$ holomorph und $\bar{B}_r(z_0) \subset U$. Für jedes $z \in B_r(z_0)$ gilt dann*

$$f(z) = \frac{1}{2\pi i} \int_{\partial B_r(z_0)} \frac{f(w)}{w-z} dw.$$

Hierbei bezeichnet $\partial B_r(z_0)$ den geschlossenen Weg $t \mapsto z_0 + re^{it}$, $t \in [0, 2\pi]$.

Beweis. Die Funktion $w \mapsto \frac{f(w)}{w-z}$ ist holomorph auf $U \setminus z$. Falls $s + |z - z_0| \leq r$, so ist $\partial B_s(z)$ ebenfalls ein geschlossener Weg in U. Nun gibt es eine Homotopie geschlossener Wege in U von $\partial B_s(z)$ nach $\partial B_r(z_0)$. Aus Korollar 6.21 ergibt sich

$$\frac{1}{2\pi i} \int_{\partial B_r(z_0)} \frac{f(w)}{w-z} dw = \frac{1}{2\pi i} \int_{\partial B_s(z)} \frac{f(w)}{w-z} dw$$

für alle kleinen $s > 0$. Die Definition des Kurvenintegrals ergibt

$$\frac{1}{2\pi i} \int_{\partial B_s(z)} \frac{f(w)}{w-z} dw = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z + se^{it})}{se^{it}} sie^{it} dt = \frac{1}{2\pi} \int_0^{2\pi} f(z + se^{it}) dt.$$

Die Funktion f ist stetig auf $\bar{B}_r(z_0)$ und daher beschränkt, sagen wir durch $C \geq 0$. Aus der Stetigkeit von f und dem Satz über dominierte Konvergenz folgt

$$\lim_{s \to 0} \frac{1}{2\pi} \int_0^{2\pi} f(z + se^{it}) dt = \frac{1}{2\pi} \int_0^{2\pi} f(z) dt = f(z),$$

was behauptet war. □

Satz 6.29 hat zahllose außerordentlich überraschende Konsequenzen. Zunächst rechne

$$\frac{1}{w-z} = \frac{1}{(w-z_0) - (z-z_0)} = \frac{1}{w-z_0} \frac{1}{1 - \frac{z-z_0}{w-z_0}} = \frac{1}{w-z_0} \sum_{k=0}^{\infty} (\frac{z-z_0}{w-z_0})^k.$$

Diese Reihe konvergiert absolut, falls $|z - z_0| < |w - z_0|$. In die Integralformel eingesetzt sehen wir für $|z - z_0| < r$:

$$f(z) = \frac{1}{2\pi i} \int_{\partial B_r(z_0)} \sum_{k=0}^{\infty} \frac{f(w)}{(w-z_0)^{k+1}} (z-z_0)^k dw.$$

Ferner ist die Konvergenz gleichmäßig auf der Menge $\{w||w-z_0|=r\}$, und wir dürfen Summation und Integration vertauschen:

$$f(z)=\sum_{k=0}^{\infty}\frac{1}{2\pi i}\int_{\partial B_r(z_0)}\frac{f(w)}{(w-z_0)^{k+1}}(z-z_0)^k dw=$$

$$=\sum_{k=0}^{\infty}(\frac{1}{2\pi i}\int_{\partial B_r(z_0)}\frac{f(w)}{(w-z_0)^{k+1}}dw)(z-z_0)^k.$$

Diese Gleichung gilt für alle z in der Kreisscheibe $B_r(z_0)$. Man werfe einen genaueren Blick auf diese Formel: der Ausdruck $\frac{1}{2\pi i}\int_{\partial B_r(z_0)}\frac{f(w)}{(w-z_0)^{k+1}}dw$ hängt nicht von z ab!! Mit anderen Worten:

Satz 6.30 (Potenzreihenentwicklung)**.** *Sei $U\subset\mathbb{C}$ offen und $f:U\to\mathbb{C}$ holomorph. Ferner sei $\bar{B}_r(z_0)\subset U$. Dann konvergiert die Potenzreihe*

$$\sum_{k=0}^{\infty}(\frac{1}{2\pi i}\int_{\partial B_r(z_0)}\frac{f(w)}{(w-z_0)^{k+1}}dw)(z-z_0)^k$$

auf der Kreisscheibe $B_r(z_0)$ gegen f. Insbesondere ist f beliebig oft komplex differenzierbar (wegen Satz 6.10), und es gilt

$$f^{(k)}(z_0)=\frac{k!}{2\pi i}\int_{\partial B_r(z_0)}\frac{f(w)}{(w-z_0)^{k+1}}dw.$$

Bemerkung 6.31. *Die erste auffallende Konsequenz aus dem Satz ist die* automatische Regularität*: die Bedingung, dass f holomorph ist, bezieht nur die ersten Ableitungen ein, es folgt aber, dass f eine glatte Funktion sein muss. Dieses Resultat hat keine Parallele in der elementaren reellen Analysis. Der "wirkliche Grund", warum diese Aussage gilt, ist dass holomorphe Funktionen Lösungen der partiellen Differentialgleichung $(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y})f=0$ sind. Diese Differentialgleichung hat die spezielle Eigenschaft,* elliptisch *zu sein, und in der* Regularitätstheorie elliptischer Differentialgleichungen *wird gezeigt, dass Lösung elliptischer Gleichungen immer glatt sind. Darauf näher einzugehen, führte zu weit, es sei aber nur gesagt, dass die Theorie der Fouriertransformation den ersten Schritt zu einem Beweis gibt.*

Der Satz gibt eine einfache Erklärung für das Verhalten mancher Potenzreihen im Reellen. Wir betrachten die Funktionen $f(z)=\frac{1}{1+z}$ und $g(z)=\frac{1}{1+z^2}$. Sie haben Potenzreihenentwicklungen

$$f(z)=\sum_{k=0}^{\infty}(-1)^k z^k$$

und

$$g(z)=\sum_{k=0}^{\infty}(-1)^k z^{2k}.$$

Beide Potenzreihen haben den Konvergenzradius 1, wie man leicht einsieht (Quotientenkriterium). Dass die Potenzreihe von f höchstens den Konvergenzradius 1 haben kann, ist klar, denn f hat bei $z=-1$ einen Pol. Die Funktion g ist aber auf ganz $\mathbb{R}$ beliebig oft stetig differenzierbar. Aber g hat eine Singularität bei $z=\pm i$, und daher kann der Konvergenzradius von g höchstens 1 sein.

Korollar 6.32 (Abschätzungen für die Ableitungen)**.** *Sei $f : B_R(z_0) \to \mathbb{C}$ holomorph und beschränkt, $|f(z)| \leq C$ für alle z. Dann gilt*

$$|f^{(k)}(z_0)| \leq \frac{Ck!}{R^k}.$$

Beweis. Für $r < R$ folgt aus Satz 6.30

$$|f^{(k)}(z_0)| = \frac{k!}{2\pi} \Big| \int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^{k+1}} dw \Big| = \frac{k!}{2\pi} \Big| \int_0^{2\pi} \frac{f(z_0 + re^{it})}{(re^{it})^{k+1}} ire^{it} dt \Big| \leq$$

$$\leq \frac{k!}{2\pi} \int_0^{2\pi} \frac{C}{r^k} dt \leq \frac{Ck!}{r^k}.$$

Grenzübergang $r \to R$ zeigt das gewünschte Ergebnis. □

Korollar 6.33 (Satz von Liouville)**.** *Sei $f : \mathbb{C} \to \mathbb{C}$ beschränkt und holomorph. Dann ist f konstant.*

Beweis. Es gilt nämlich nach Korollar 6.32, falls $|f(z)| \leq C$, $R > 0$

$$|f^{(k)}(0)| \leq \frac{k!}{2\pi} C \frac{1}{R^k}$$

für alle R, also sind alle Ableitungen $f^{(k)}(0) = 0$ $(k \geq 1)$. Daher verschwinden alle Koeffizienten in der Potenzreihenentwicklung von f außer dem 0ten. □

Dieses Ergebnis erlaubt einen eleganten Beweis des Fundamentalsatzes der Algebra.

Satz 6.34 (Fundamentalsatz der Algebra)**.** *Sei $f : \mathbb{C} \to \mathbb{C}$ ein nichtkonstantes Polynom mit komplexen Koeffizienten. Dann hat f eine Nullstelle.*

Beweis. Sei $f(z) = a_n z^n + q(z)$, wobei $a_n \neq 0$ und q ein Polynom vom Grad $\leq n-1$ sei. Es gibt dann $R > 0$ so dass für $|z| \geq R$ gilt $|f(z)| \geq \frac{1}{2}|a_n||z^n|$. Wenn f keine Nullstelle hätte, dann folgt dass $1/f$ holomorph auf ganz $\mathbb{C}$ und beschränkt ist. Somit ist $1/f$ konstant. □

Es scheint als seien $\sin(z)$ und $\cos(z)$ Gegenbeispiele zum Satz von Liouville. Es gilt aber

$$\cos(ix) = \frac{1}{2}(e^{i^2 x} + e^{-i^2 x}) = -\frac{1}{2}(e^x + e^{-x}),$$

also ist der Cosinus nicht beschränkt.

Ein weiterer Aspekt, der im Reellen keine Parallele hat, ist der

Satz 6.35 (Riemannscher Hebbarkeitssatz)**.** *Es sei $f : U \setminus z_0 \to \mathbb{C}$ holomorph und $z_0 \in U$. Angenommen, es gibt $R > 0$ so dass $B_R(z_0) \subset U$ und so dass $f|_{B_R(z_0)}$ beschränkt ist. Dann lässt sich f holomorph auf ganz U fortsetzen (natürlich ist die Fortsetzung eindeutig bestimmt).*

Man könnte meinen, die Funktion $\cos(1/z)$ sei ein Gegenbeispiel zum Hebbarkeitssatz, aber das ist nicht beschränkt, wenn man ins Komplexe geht, denn $\cos(1/ix) = \frac{1}{2}(e^{1/x} + e^{-1/x})$ ist für $x \to 0$ unbeschränkt.

Beweis. Ohne Beschränkung der Allgemeinheit ist $z_0 = 0$. Sei $|f(z)| \leq C$ für alle $0 < |z| < R$. Setze

$$h(z) := \begin{cases} z^2 f(z) & z \neq 0 \\ 0 & z = 0. \end{cases}$$

Wenn wir zeigen können, dass h holomorph mit $h'(0) = 0$ ist, dann können wir h in eine Potenzreihe entwickeln

$$h(z) = \sum_{k=2}^{\infty} a_k z^k$$

und $\sum_{k=0}^{\infty} a_{k+2} z^k$ ist die gesuchte Fortsetzung von f. Offenbar ist h stetig, und es gilt

$$\left|\frac{h(z) - h(0)}{z}\right| \le \frac{|z|^2 C}{|z|} \to 0$$

und daher ist h differenzierbar in 0 mit $h'(0) = 0$. Es fehlt das Argument, dass h' *stetig* in 0 ist, und hier ist der wesentliche Punkt. Es ist $h'(z) = 2zf(z) + z^2 f'(z)$ für $z \neq 0$. Nun folgt aus den Abschätzungen für die Ableitung 6.32

$$|f'(z)| \le \frac{1}{|z|} C$$

und daher

$$|2zf(z) + z^2 f'(z)| \le 2|z|C + \frac{|z|^2}{|z|} C = 3C|z|$$

und daher ist h' stetig in 0. □

Es gibt einen bedeutsamen Zusammenhang von Potenzreihen mit Fourierreihen.

Proposition 6.36. *Es sei $U \subset \mathbb{C}$ offen, $f : U \to \mathbb{C}$ holomorph und $\bar{B}_r(z_0) \subset U$. Betrachte die 2π-periodische Funktion $g(t) = \frac{1}{\sqrt{2\pi}} f(z_0 + re^{it})$. Dann gilt für die Fourierkoeffizienten von g:*

$$\hat{g}(k) = \begin{cases} r^k \frac{1}{k!} f^{(k)}(z_0) & k \ge 0 \\ 0 & k < 0. \end{cases}$$

Beweis. Es gilt

$$\frac{1}{2\pi i} \int_{\partial B_r(z_0)} \frac{f(z)}{(z - z_0)^{k+1}} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{it})}{r^{k+1} e^{i(k+1)t}} ire^{it} dt =$$

$$= \frac{1}{2\pi r^k} \int_0^{2\pi} f(z_0 + re^{it}) e^{-ikt} dt = \frac{1}{r^k} \hat{g}(k).$$

Falls $k < 0$, so ist die Funktion $\frac{f(z)}{(z-z_0)^{k+1}}$ auf ganz $\bar{B}_r(z_0)$ holomorph und daher gilt $\hat{g}(k) = 0$ nach dem Cauchy'schen Integralsatz. Für $k \ge 0$ ergibt sich aus den Formeln für die Ableitungen

$$\hat{g}(k) = r^k \frac{1}{k!} f^{(k)}(z_0).$$ □

Mit der Parzeval-Gleichung 5.24 folgt

$$\sum_{k=0}^{\infty} r^{2k} \left|\frac{f^{(k)}(z_0)}{k!}\right|^2 = \|g\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{it})|^2 dt, \tag{6.37}$$

was wiederum eine interessante Konsequenz hat.

Satz 6.38 (Lokales Maximumsprinzip)**.** *Es sei $f : B_R(z_0) \to \mathbb{C}$ holomorph. Falls f in z_0 ein lokales Maximum besitzt, so ist f konstant.*

Beweis. Falls f ein lokales Maximum in z_0 hat, so gilt $|f(z_0)| \geq |f(z)|$ für alle $z \in B_R(z_0)$. Sei $0 < r < R$. Aus 6.37 folgt

$$\sum_{k=0}^{\infty} r^{2k} |\frac{f^{(k)}(z_0)}{k!}|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{it})|^2 dt \leq |f(z_0)|^2.$$

Dies ist nur möglich, wenn $f^{(k)}(z_0) = 0$ für alle $k > 0$. Dann ist wegen der Potenzreihenentwicklung f konstant. $\square$

Sei $f : U \to \mathbb{C}$ holomorph und $\bar{B}_R(z_0) \subset U$. Dann sind nach der Cauchy'schen Integralformel die Werte von f in der Kreisscheibe $B_R(z_0)$ durch das Verhalten von f auf dem Rand der Kreisscheibe bestimmt. Dieses Phänomen (ebenfalls ohne Parallele im Reellen) werden wir erheblich verschärfen:

Lemma 6.39. *Sei $f : U \to \mathbb{C}$ holomorph und $f(z_0) = 0$. Entweder ist f konstant auf einer Kreisscheibe um z_0, oder es gibt $r > 0$, so dass es keine weitere Nullstelle z von f mit $|z - z_0| < r$ gibt.*

Beweis. Betrachte die Potenzreihenentwicklung von f um z_0: $f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$. Nach Voraussetzung ist $a_0 = 0$. Wenn f auf der Kreisscheibe nicht konstant Null ist, so gibt es ein kleinstes n mit $a_n \neq 0$. Dann können wir f als $f(z) = (z - z_0)^n g(z)$ mit einer holomorphen Funktion g schreiben. Da $g(z_0) \neq 0$ und weil g stetig ist, hat g keine Nullstellen nahe z_0, und dann f auch nicht. $\square$

Satz 6.40 (Identitätssatz)**.** *Sei U wegzusammenhängend und $f, g : U \to \mathbb{C}$ holomorph. Falls eine Folge $z_n \to z_0 \in U$ existiert mit $z_n \neq z_0$ und $f(z_n) = g(z_n)$, so gilt $f = g$ auf ganz U.*

Beweis. Wir dürfen $g = 0$ annehmen. Aus Stetigkeitsgründen folgt $f(z_0) = 0$. Nach Lemma 6.39 ist f also konstant 0 auf einer Kreisscheibe $B_r(z_0)$. Dieses Argument "globalisiert" man wie folgt. Es sei $z \in U$ und $c : [0,1] \to U$ ein Weg von z_0 nach z. Sei $s := \sup\{t | f \circ c|_{[0,t]} \equiv 0\}$. Wir behaupten, dass $s = 1$ ist. Wir argumentieren durch Widerspruch und nehmen $s < 1$ an. Falls $c|_{[0,s]}$ konstant ist, so ist $c(s) = z_0$, und es gibt dann ein $\delta > 0$ mit $f(c(s+u)) = 0$ für $u < \delta$ (man wählt nämlich δ so klein, dass $c(s+u) \in B_r(z_0)$ für $u < \delta$, und wir wissen schon, dass f auf $B_r(z_0)$ identisch Null ist. Dies steht im Widerspruch zur Definition von s. Wenn $c|_{[0,s]}$ nicht konstant ist, so gibt es eine Folge $s_n < s$, $c(s_n) \neq c(s)$ (denn dann nimmt $c| - [0,s]$ unendlich viele Werte an). Weil $f(c(s_n)) = 0$, muss $f(c(s)) = 0$ gelten, und nach Lemma 6.39 ist dann $f \equiv 0$ auf $B_{r'}(c(s))$. Mit demselben Argument wie eben findet man $\delta > 0$, so dass $f \circ c|_{[0,\delta]} \equiv 0$, ein Widerspruch.[1] $\square$

Satz 6.41 (Globales Maximumsprinzip)**.** *Sei $U \subset \mathbb{C}$ wegzusammenhängend und $f : U \to \mathbb{C}$. Falls $|f|$ in $z_0 \in U$ ein lokales Maximum annimmt, so ist f konstant.*

Beweis. Nach dem lokalen Maximumsprinzip ist f konstant auf einer Kreisscheibe um z_0 und daher nach dem Identitätssatz konstant auf ganz U. $\square$

[1] Der in der Vorlesung gegebene Beweis war nicht ganz vollständig.

6.7. **Die Windungszahl.** Es folgt nun ein erster Einblick in die algebraische Topologie.

Definition 6.42. *Sei $a \in \mathbb{C}$ und c ein geschlossener C^2-Weg in $\mathbb{C} \setminus a$. Die* Windungszahl *von c um a ist*

$$\deg(c,a) := \frac{1}{2\pi i} \int_c \frac{dz}{z-a}.$$

Beispiel: sei $\gamma(t) = e^{ikt}$. Dann gilt $\deg(\gamma, z) = k$, falls $|z| < 1$ und 0 sonst. A priori ist die Windungsszahl eine komplexe Zahl. Wir wollen nun zeigen, dass dies eine ganze Zahl ist.

Lemma 6.43. *Die Windungszahl ist eine ganze Zahl.*

Beweis. Die Aussage $\deg(c,a) \in \mathbb{Z}$ ist äquivalent zu $\exp(2\pi i \deg(c,a)) = 1$. Die Definition des Kurvenintegrals zeigt

$$\exp(2\pi i \deg(c,a)) = \exp(\int_c \frac{dz}{z-a}) = \exp(-\int_0^1 \frac{c'(t)}{c(t)-a} dt)$$

Setze für $s \in [0,1]$:

$$G(s) := (c(s)-a) \exp(-\int_0^s \frac{c'(t)}{c(t)-a} dt).$$

Es gilt $G(0) = (c(0)-a)$ und $c(s) - a \neq 0$. Differenzieren zeigt

$$G'(s) = \frac{c'(s)}{c(s)-a} G(s) - \frac{c'(s)}{c(s)-a} G(s) = 0.$$

Also ist G konstant, das heißt

$$(c(0)-a) = G(0) = G(1) = (c(1)-a) e^{-2\pi i \deg(c,a)}.$$

weil $c(0) = c(1) \neq a$ folgt $e^{-2\pi i \deg(c,a)} = 1$, also $\deg(c,a) \in \mathbb{Z}$. □

Lemma 6.44. *Seien c_0, c_1 zwei geschlossene C^2-Wege in $\mathbb{C} \setminus \{a\}$. Angenommen, es gebe eine Homotopie von geschlossenen Wegen in $\mathbb{C} \setminus \{a\}$ zwischen c_0 und c_1. Dann gilt $\deg(c_0, a) = \deg(c_1, a)$.*

Beweis. Folgt sofort aus Satz 6.27. □

Lemma 6.45. *Es sei $c : [0,1] \to \mathbb{C}$ ein geschlossener C^1-Weg. Dann ist die Funktion $\mathbb{C} \setminus c([0,1]) \to \mathbb{Z}$, $a \mapsto \deg(c,a)$ lokal-konstant.*

Beweis. Es seien $a, a_n \in \mathbb{C} \setminus c([0,1])$ und $a_n \to a$. Es reicht, zu zeigen, dass $\deg(c, a_n) \to \deg(c,a)$, denn aufgrund der Ganzzahligkeit der Windungszahl folgt dann $\deg(c,a_n) = \deg(c,a)$ für genügend große n. Nun gilt

$$\deg(c,a_n) - \deg(c,a) = \frac{1}{2\pi i} \int_c \frac{1}{z-a_n} - \frac{1}{z-a} dz =$$

$$= \frac{1}{2\pi i} \int_c \frac{a_n - a}{(z-a_n)(z-a)} dz = \frac{1}{2\pi i} \int_0^1 \frac{a_n - a}{(c(t)-a_n)(c(t)-a)} c'(t) dt.$$

Weil $K = \{a\} \cup \{a_n | n \in \mathbb{N}\}$ kompakt ist, ist $\mathrm{dist}(K, c([0,1])) = \delta > 0$. Der Nenner im Integral ist also nach unten durch δ^2 beschränkt. Somit findet man leicht eine integrierbare Majorante, und weil der Integrand gegen Null konvergiert, folgt $\deg(c,a_n) \to \deg(c,a)$ aus dem Satz über dominierte Konvergenz. □

Satz 6.46. *Es sei $a \in \mathbb{C}$ und c_0, c_1 zwei geschlossene Wege in $\mathbb{C} \backslash a$. Falls $\deg(c_0, a) = \deg(c_1, a)$, so gibt es eine Homotopie geschlossener Wege von c_0 nach c_1 in $\mathbb{C} \setminus a$.*

Beweis. Sei zuerst $c : [0,1] \to \mathbb{C}^{\backslash}$ ein geschlossener Weg mit $c(0) = c(1) = 1$ und $\deg(c, 0) = 0$. Setze

$$g_s(x) := s \int_0^x \frac{c'(t)}{c(t)} dt$$

und

$$h(s,t) := \exp(s \int_0^t \frac{c'(x)}{c(x)} dx).$$

Es gilt dann $h(0,t) = 1$, $h(s,0) = 1$ sowie $h(s,t) \in \mathbb{C}^{\backslash}$. Der wesentliche Punkt ist, dass

$$h(s,1) = \exp(s \int_0^1 \frac{c'(x)}{c(x)} dx) = \exp(s 2\pi i \deg(c,0)) = \exp(0) = 1 = h(s,0)$$

gilt. Deshalb ist $h(s,t)$ eine Homotopie geschlossener Wege. Schließlich gilt

$$c(t) h(1,t)^{-1} = c(t) \exp(-1 \int_0^t \frac{c'(x)}{c(x)} dx)$$

und Differentation der rechten Seite zeigt

$$\frac{d}{dt}(c(t)h(1,t)^{-1}) = (c'(t) - c(t)\frac{c'(x)}{c(x)}) \exp(-1 \int_0^t \frac{c'(x)}{c(x)} dx) = 0,$$

also gilt $c(t) = h(1,t)$. Somit haben wir den Satz in dem Spezialfall $a = 0$, $c_0(0) = c_0(1) = 1$, $c_1(t) \equiv 1$ bewiesen. Den allgemeinen Fall führt man darauf zurück, in mehreren Schritten.

Falls $a = 0$, $c(0) = c(1) \neq 1$ und $\deg(c, 0) = 0$, so wählen wir wir einen Weg $b : [0,1] \to \mathbb{C}^\times$ mit $b(0) = 1$, $b(1) = c(0)^{-1}$ und betrachten die Homotopie $h(s,t) := b(s)c(t)$. Dies ist eine Homotopie geschlossener Wege in $\mathbb{C}^\times$, $h_0(t) = c(t)$ und $h_1(0) = h_1(1) = 1$. Dies führt die Frage auf den Fall $c(0) = 1$ zurück.

Für zwei geschlossene Wege $c_0, c_1 : [0,1] \to \mathbb{C}^\times$ dürfen wir das Produkt $c_0 c_1$ und den Kehrwert c_0^{-1} nehmen, und das sind wieder geschlossene Wege. Man rechnet

$$\deg(c_0 c_1, 0) = \frac{1}{2\pi i} \int_0^1 \frac{c_0'(t)c_1(t) + c_0(t)c_1'(t)}{c_0(t)c_1(t)} dt = \deg(c_0, 0) + \deg(c_1, 0)$$

und

$$\deg(c_0^{-1}, 0) = \frac{1}{2\pi i} \int_0^1 \frac{-c_0'(t)}{c_0(t)^2 c_0(t)^{-1}} dt = -\deg(c_0, 0).$$

Sind also c_0 und c_1 zwei geschlossene Wege in $\mathbb{C}^\times$ mit $\deg(c_0, 0) = \deg(c_1, 0)$, so gilt $\deg(\frac{c_0}{c_1}, 0) = 0$. Also ist der geschlossene Weg $\frac{c_0}{c_1}$ homotop zu dem konstanten Weg bei 1, und sei $h(s,t)$ eine Nullhomotopie von $\frac{c_0}{c_1}$. Dann ist $h(s,t)c_1(t)$ eine Homotopie zwischen c_0 und c_1.

Der letzte Schritt ist die Verallgemeinerung auf den Fall $a \neq 0$. Die Rechnung

$$\deg(c, a) = \frac{1}{2\pi i} \int_0^1 \frac{c'(t)}{c(t) - a} dt = \frac{1}{2\pi i} \int_0^1 \frac{c'(t) - a}{c(t) - a} dt = \deg(c - a, 0)$$

aber zeigt, dass die Windungszahl des verschobenen Weges $t \mapsto c(t) - a$ um 0 gleich $\deg(c, a)$ ist. Verschiebung um a beendet dann den Beweis. $\square$

6.8. **Der Residuensatz.**

Definition 6.47. *Sei $U \subset \mathbb{C}$ offen und $z_0 \in U$. Sei $f : U \setminus z_0 \to \mathbb{C}$ holomorph. Wir sagen, dass f* eine Polstelle in z_0 hat*, falls $k \in \mathbb{N}_0$ und ein $r > 0$ existiert und $C \geq 0$, so dass $|f(z)(z - z_0)^k| \leq C$ für alle $0 < |z - z_0| < r$ gilt. Das kleinste k mit dieser Eigenschaft heißt* Ordnung der Polstelle z_0 von f.

Falls $g : U \to \mathbb{C}$ holomorph ist, so hat $\frac{g(z)}{(z-z_0)^k}$ einen Pol der Ordnung $\leq k$ in z_0, und wenn $g(z_0) \neq 0$, so hat dieser Pol die Ordnung k.

Der Hebbarkeitssatz besagt, dass, falls f in z_0 einen Pol der Ordnung 0 hat, die Funktion f holomorph auf $U \cup \{z_0\}$ fortgesetzt werden kann. Dieser Schluss lässt sich verallgemeinern.

Satz 6.48. *$f : U \setminus z_0 \to \mathbb{C}$ hat genau dann einen Pol der Ordnung k bei z_0, wenn eine holomorphe Funktion $g : U \to \mathbb{C}$ existiert mit $g(z_0) \neq 0$ und $\frac{g(z)}{(z-z_0)^k} = f(z)$.*

Beweis. Wenn f einen Pol der Ordnung $\leq k$ in z_0 hat, so ist die Funktion $g(z) := f(z)(z - z_0)^k$ beschränkt in einer Umgebung von z_0, besitzt also nach dem Riemannschen Hebbarkeitssatz eine holomorhe Fortsetzung, welche ebenfalls mit g bezeichnet werde. Sei $g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ die Potenzreihenentwicklung von g um z_0. Wenn die Ordnung der Polstelle genau k ist, so ist

$$f(z)(z - z_0)^{k-1} = \frac{g(z)}{z - z_0} = \sum_{n=0}^{\infty} a_n(z - z_0)^{n-1} = \frac{a_0}{z - z_0} + \sum_{n=1}^{\infty} a_n(z - z_0)^{n-1}$$

nicht beschränkt um z_0. Weil der zweite Summand holomorph ist, muss dann $a_0 \neq 0$ sein, also $g(z_0) = a_0 \neq 0$. □

In der Situation des vorherigen Satzes schreiben wir

$$f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^{n-k} = \sum_{n=0}^{k-1} a_n \frac{1}{(z - z_0)^{k-n}} + \sum_{n=k}^{\infty} a_n(z-z_0)^{n-k} =: H_{z_0} f(z) + h(z).$$

Der zweite Summand h ist holomorph in ganz U und der erste Summand heißt *Hauptteil von f bei z_0*. Für genügend kleine $r > 0$ (d.h. $\bar{B}_r(z_0) \subset U$) rechnen wir

$$\int_{\partial B_r(z_0)} f(z)dz = \int_{\partial B_r(z_0)} \sum_{n=0}^{k-1} a_n \frac{1}{(z - z_0)^{k-n}} dz + \int_{\partial B_r(z_0)} h(z)dz =$$

$$\int_{\partial B_r(z_0)} a_{k-1} dz = 2\pi i a_{k-1}.$$

Das Kurvenintegral von h ist Null wegen des Cauchy'schen Integralsatzes.

Definition 6.49. *Es sei $U \subset \mathbb{C}$ offen und $f : U \setminus z_0 \to \mathbb{C}$ sei holomorph und habe eine Polstelle in z_0. Die Zahl*

$$\operatorname{res}(f, z_0) := \frac{1}{2\pi i} \int_{\partial B_r(z_0)} f(z)dz,$$

wobei $r > 0$ so klein sein muss, dass $\bar{B}_r(z_0) \subset U$ gilt, heißt Residuum *von f bei z_0.*

Nun formulieren wir eine globale Version dieser Begriffe.

Definition 6.50. *Eine Teilmenge $S \subset U$ heißt* diskret*, falls zu jedem $z \in S$ ein $r > 0$ existiert, so dass $B_r(z) \cap S = \{z\}$ gilt.*

Beispielsweise sind Nullstellenmengen holomorpher Funktionen (die nicht identisch Null sind) diskret, nach dem Identitätssatz. Betrachte etwa die Funktion $f(z) = \sin(\pi \frac{1}{z})$, $f : \mathbb{C}^\times \to \mathbb{C}$. Die Nullstellenmenge ist $S = \{\frac{1}{n} | 0 \neq n \in \mathbb{Z}\}$, und das ist diskret in $\mathbb{C}^\times$ (aber natürlich nicht in ganz $\mathbb{C}$). Man beachte, dass falls $S \subset U$ diskret ist, auch jede Teilmenge $T \subset U$ diskret ist. Ferner sind diskrete Mengen abgeschlossen in U.

Definition 6.51. *Sei $U \subset \mathbb{C}$ offen und $S \subset U$ diskret. Eine holomorphe Funktion $f : U \setminus S \to \mathbb{C}$ heißt* meromorph *auf U, falls f in jedem $z \in S$ einen Pol hat. Die Menge S heißt* Singularitätenmenge *von f.*

Man beachte, dass meromorphe Funktionen auf U nicht auf ganz U definiert sein müssen (nämlich nicht in den Polen). Rationale Funktionen sind meromorph auf ganz $\mathbb{C}$. Hingegen ist der komplexe Logarithmus *nicht* meromorph auf $\mathbb{C}$, denn er ist auf der ganzen negativen reellen Achse $\{z | z \leq 0\}$ nicht definiert, und diese Menge ist nicht diskret in $\mathbb{C}$. Wir kommen nun zum Ziel der Vorlesung.

Satz 6.52 (Residuensatz)**.** *Sei $U \subset \mathbb{C}$ offen und f eine meromorphe Funktion auf U, mit Singularitätenmenge S. Sei $c : [a, b] \to U$ ein geschlossener Weg, so dass $c([a, b]) \cap S = \emptyset$ und so dass c (in U!!) nullhomotop ist. Unter diesen Voraussetzungen gilt*

$$\frac{1}{2\pi i} \int_c f(z)dz = \sum_{w \in S} \deg(c, w) \operatorname{res}(f, w),$$

und die Summe auf der rechten Seite ist endlich (also $\deg(c, w) = 0$ für alle bis auf endlich viele $w \in S$).

Beweis. Sei $h(s, t)$ eine Nullhomotopie von c in U. Das Bild $K = h([0, 1] \times [a, b]) \subset U$ ist kompakt. Es folgt, dass $K \cap S$ endlich ist. Um dies einzusehen, wählen wir für $w \in S$ ein $r_w > 0$ mit $B_{r_w}(w) \subset U$ und $B_{r_w}(w) \cap S = \{w\}$ und betrachten die offene Überdeckung

$$\mathcal{U} = \{U \setminus S\} \cup \{B_{r_w}(w) | w \in S\}$$

von U. Diese Überdeckung ist so gewählt, dass jedes $w \in S$ in genau einer der überdeckenden Mengen liegt. Nach dem Satz von Heine-Borel gibt es endlich viele $w_1, \ldots, w_m$, so dass

$$K \subset (U \setminus S) \cup B_{r_{w_1}}(w_1) \cup \ldots \cup B_{r_{w_m}}(w_m).$$

Damit ist $K \cap S = \{w_1, \ldots, w_m\}$ endlich. Setze $T := S \setminus \{w_1, \ldots, w_m\}$ und $U_0 := U \setminus T$. Nach Konstruktion ist U_0 offen in $\mathbb{C}$, f ist meromorph auf U mit *endlich vielen Polstellen.* Ferner ist c ein geschlossener Weg in U_0, und *nullhomotop in* U_0.

Ist $w \in T$, so ist c ein nullhomotoper Weg in $\mathbb{C} \setminus \{w\}$, folglich gilt

$$\deg(c, w) = 0 \forall w \in T. \tag{6.53}$$

Dies zeigt, dass die Summe auf der rechten Seite endlich ist. Wir schränken uns auf U_0 ein. Seien $g_1, \ldots, g_m$ die Hauptteile der Funktion f bei den Polstellen $w_1, \ldots, w_m$. Dann ist die Funktion $h(z) := f(z) - \sum_{j=1}^m g_j(z)$ holomorph in U_0. Nun rechne

$$\frac{1}{2\pi i} \int_c f(z)dz = \frac{1}{2\pi i} \int_c h(z)dz + \frac{1}{2\pi i} \int_c \sum_{j=1}^m g_j(z)dz. \tag{6.54}$$

Weil c nullhomotop in U_0 ist, zeigt der Cauchy-Integralsatz, dass $\frac{1}{2\pi i}\int_c h(z)dz = 0$. Nun können wir den Hauptteil $g_j(z)$ als

$$\sum_{n=1}^{k_j} a_n \frac{1}{(z-w_j)^n}$$

schreiben. Durch Integration um den Kreis $\partial B_r(w_j)$ sehen wir, dass $\operatorname{res}(f,w_j) = \operatorname{res}(g_j,w_j) = a_1$ gilt. Es folgt

$$\frac{1}{2\pi i}\int_c g_j(z)dz = \sum_{n=1}^{k_j} \frac{1}{2\pi i}\int_c a_n \frac{1}{(z-w_j)^n}dz = \frac{1}{2\pi i}\int_c a_1 \frac{1}{(z-w_j)^1}dz, \tag{6.55}$$

weil die Funktionen $\frac{1}{(z-w_j)^n}$ für $n \geq 2$ Stammfunktionen besitzen. Schließlich ist nach Definition der Windungszzahl

$$\frac{1}{2\pi i}\int_c a_1 \frac{1}{(z-w_j)^1}dz = \deg(c,w_j)\operatorname{res}(f,w_j). \tag{6.56}$$

Die Formeln 6.53, 6.54, 6.55 und 6.56 zeigen dann die behauptete Integralformel. □

Literatur

[1] T. Bröcker: *Analysis II.*
[2] E. M. Stein, R. Shakarchi: *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press